

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-251101

(43)Date of publication of application : 27.09.1996

(51)Int.Cl. H04B 10/02
H04B 10/20

(21)Application number : 07-279365

(71)Applicant : HEWLETT PACKARD CO <HP>

(22)Date of filing : 26.10.1995

(72)Inventor : GRANT ROBERT H
STOEVHASE BENT
PUROHIT ROBIN
SULLIVAN GREGORY T
BOOK DAVID

(30)Priority

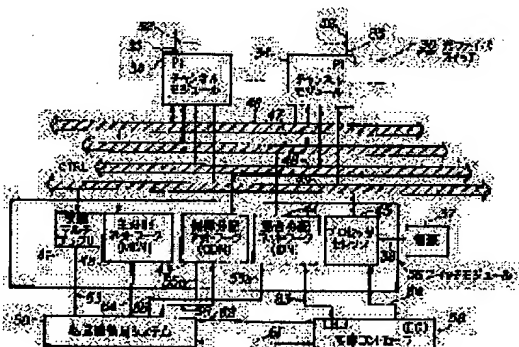
Priority number : 94 330273 Priority date : 27.10.1994 Priority country : US

(54) TRANSFER PATH ALLOCATION SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To obtain high performance and flexibility for various applications, to guarantee that a transfer request is processed finally by a switch without depending upon the probability that a port is usable, and to enable both circuit switching and frame switching.

SOLUTION: The transfer path allocation system 50 is provided with ports 33 corresponding to respective channels 32, a memory means which is related to the ports 33, a distribution network, a scheduler, and an arbitrator which is so constituted as to reject and allow a transfer request. Then a source fairness guaranteeing means includes at least one destination port 33 which is related to a source port 33 for determining the order of future access to the source port 33 as the source of data and has its order corresponding to a rejected transfer request.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2000 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-251101

(43) 公開日 平成8年(1996)9月27日

(51) Int.Cl.⁹

H 0 4 B 10/02
10/20

識別記号

庁内整理番号

F I

H 0 4 B 9/00

技術表示箇所

T
N

審査請求 未請求 請求項の数 1 O L (全 23 頁)

(21) 出願番号 特願平7-279365

(22) 出願日 平成7年(1995)10月26日

(31) 優先権主張番号 330-273

(32) 優先日 1994年10月27日

(33) 優先権主張国 米国 (U S)

(71) 出願人 590000400

ヒューレット・パッカード・カンパニー
アメリカ合衆国カリフォルニア州パロアル
ト ハノーバー・ストリート 3000

(72) 発明者 ロバート・エイチ・グラント

カナダ国オンタリオ州トロント デアポー
ン・アヴェニュー 108

(72) 発明者 ベント・ストーヴベイス

カナダ国オンタリオ州トロント カールト
ン・ストリート 20、アパートメント
1224

(74) 代理人 弁理士 荻野 平 (外 5 名)

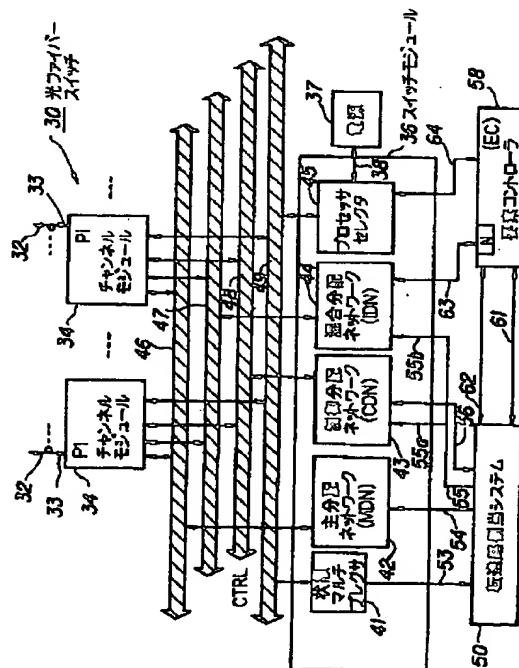
最終頁に続く

(54) 【発明の名称】 転送路割り当てシステム

(57) 【要約】

【課題】 高性能で、さまざまなアプリケーションに対するフレキシビリティがあり、転送要求がポートの利用可能である確率に依存することなく最終的にスイッチによって処理されることを保証でき、回路切り換えとフレーム切り換えの両方を行なう。

【解決手段】 転送路割り当てシステム50は、チャンネル32のそれぞれに対応する複数のポート33と、ポート33に係属付けられたメモリ手段84と、分配ネットワーク362と、スケジューラ118と、転送要求を拒絶および許可するように構成されたアービトラータ123と、データ11の出所としての出所ポート33への将来のアクセスの順序を決めるための出所ポート33に係属付けられ、順序は拒絶された転送要求に対応する少なくとも1つの宛先ポート33を含む出所公平性保証手段130とからなる。



【特許請求の範囲】

【請求項 1】 光ファイバーネットワークにおいて光ファイバーチャンネル（32）を選択的に相互接続する光ファイバースイッチ（30）を介した高性能のデータ転送のための転送路割り当てシステム（50）であって、前記チャンネル（32）のそれぞれに対応する複数のポート（33）と、
出所ポート（33）から着信した新しいデータフレーム（11）を格納するための前記ポート（33）に関係付けられたメモリ手段（84）と、
前記ポート（33）を相互接続するように構成された分配ネットワーク（362）と、
前記メモリ手段（84）が前記新しいデータフレーム（11）を受け取ったことを判定するように構成された、前記メモリ手段（84）につながったセントリ（104）と、
前記ポート（33）のそれぞれに対応する宛先待ち行列（121）を維持するため、前記セントリ（104）が前記新しいデータフレーム（11）の受信を判定した後前記セントリ（104）から前記新しいデータフレーム（11）に関する、出所ポート（33）を同定し、前記メモリ手段（84）内における前記フレームの位置を同定するメモリアドレスを同定する出所ポート標識と宛先ポート（33）を同定する宛先ポート標識を含む転送路データを受け取るように構成され、また前記転送路データを前記宛先ポート（33）に対応する宛先待ち行列（121）に入れ、前記宛先待ち行列（121）から前記転送路データを検索するように構成されたスケジューラ（118）と、
前記分配ネットワーク（36）を制御するように構成され、前記スケジューラ（118）と前記ポート（33）につながれ、前記宛先ポート（33）が他のフレーム転送要求をサービスするためにビジーであること、および前記宛先ポート（33）が前記フレーム転送要求を受けることができることを判定するように構成され、また利用可能性に基づいて前記転送要求を拒絶および許可するように構成されたアービトラータ（123）と、
データ（11）の出所としての前記出所ポート（33）への将来のアクセスの順序を決めるための前記出所ポート（33）に関係付けられ、前記順序は拒絶された転送要求に対応する少なくとも1つの宛先ポート（33）を含む出所公平性保証手段（130）と、からなることを特徴とする転送路割り当てシステム（50）。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、広義にはデータ通信と光ファイバーに関し、より詳細には、すべてのデータ転送要求が最終的にポートの利用可能性に左右されることなく光ファイバースイッチによってサービスされるようにする公平性保証機構を有する、光ファイバーネッ

トワークのための光ファイバースイッチを介して光ファイバーポート間でデータ転送路を割り当てるためのフレキシブルで高性能の転送路割り当てシステムとその方法に関する。

05 【0002】

【従来の技術】データ通信ネットワークは一般に、例えばコンピュータ、周辺機器その他の要素あるいは装置間での相互通信を可能にする相互接続された通信チャンネル群を含んでいる。従来、ネットワークは同軸ケーブル構成やツイストペアケーブル構成を用いて形成され、適

10 当なインターフェースあるいはネットワークスイッチを介して相互接続された通信チャンネルを用いて構成されていた。

【0003】ネットワーク業界において、光ファイバーケーブルは、その帯域幅の広さ、良好な伝播特性およびその他の伝送特性の最適性から、同軸ケーブルやツイストペアに代わってますます広く用いられるようになって

30 いる。近年、ファイバーチャンネルプロトコルが開発され、情報システムに関する米国標準規格（ANSI）として採用された。ファイバーチャンネルの工業規格はたとえば情報システムに関する米国標準規格（ANSI）（1993年）のFibre Channel Physical And Signalling Interface, rev. 4.2に詳細に規定されている。ファイバーチャンネルの工業規格によれば、この規格に適合した光ファイバーネットワークを介して可変長のデータフレームの通信を可能にすることによって、従来の工業規格に比べて遥かに高い性能とフレキシビリティを得ることができる。

【0004】図9は、可変長のフレーム11を示す。この可変長フレーム11は、可変長フレーム11の始まりを示す2進シーケンスである4バイトのフレーム始点（SOF）標識12を有する。SOF標識12の後には、24バイトのヘッダ14が設けられる。このヘッダ14は、可変長フレーム11の出所アドレス及び宛先アドレスと、可変長フレーム11が制御情報であるか実データであるかを指定するものである。ヘッダ14の後には、可変長データ16のフィールドがある。データ16の長さは、0から2,112バイトまでである。データ16の後には、誤り検出のための4バイトのCRC（周期冗長検査）符号17があり、さらに4バイトのフレーム終点（EOF）標識18がある。図9の可変長フレーム11は、固定フレームよりはるかにフレキシブルであり、アプリケーションに応じたニーズに応えることができる点ではるかに高い性能をもたらすものである。

【0005】ファイバーチャンネルの工業規格はまた、いくつかの異なる種類のデータ転送を可能とするものである。クラス1転送には回路切り換え、すなわち、ネットワークスイッチを介した指定のデータ転送路を必要とし、またネットワーク要素間での2つ以上、しばしば多

数のデータフレームの転送が行なわれる。これに対して、クラス 2 転送では 1 つのネットワーク要素から他のネットワーク要素に 1 つのフレームが転送されるたびに、ネットワークスイッチを介した 1 つの転送路の割り当てを必要とする。

【0006】

【発明が解決しようとする課題】ファイバーチャネル工業規格に従ったネットワークを実施するための光ファイバースイッチは、まだ初期段階にある。当業界で知られるこの光ファイバースイッチの 1 つに、米国 IBM 社の製造販売する ANCHOR がある。しかし、ANCHOR インターフェースの性能は、多くのアプリケーションにおいて最適とはいえず、大幅な改善の余地がある。さらに、この不完全な ANCHOR インターフェースは、主としてクラス 1 転送の回路切り換えを可能として、クラス 2 転送のためのフレーム切り換えについては、非常に制約の大きいものであるという点でフレキシビリティに欠けるものである。

【0007】さらに、従来のほとんどの光ファイバースイッチは、利用可能なポートを割り当てるというものであった。スイッチを介した転送要求が、最終的にスイッチによってサービスされることを保証する機構は存在しなかった。かかるスイッチは確率に依存するものである。すなわち、転送要求が長期間にわたってあるポートへのアクセスを一貫して拒否されることはないと考えられる。しかし、かかる従来のスイッチでは、所望のポートへのアクセスを一貫して拒否されるデータフレームも存在する。さらに、この問題はポートアクセスのプロトコルにおける共鳴すなわち繰り返しパターンによって深刻になる、すなわち、長いシーケンスによって、1 つあるいはそれ以上のポートへの非常に規則的な周期的なトラヒックパターンが生じる可能性がある。このパターンがポート資源の割り当ての態様と一致すると、このシーケンスはポート資源の割り当てを停止し、他のシーケンスの速度を低下させる可能性さえある。

【0008】したがって、当業界において光ファイバーネットワークのためのファイバーチャネル規格を、既存のシステムよりはるかに高い性能とフレキシビリティを実現しながら実施するための、新たな改良されたシステムが必要とされている。特に、高性能で、さまざまなアプリケーションに対するフレキシビリティがあり、転送要求がポートの利用可能である確率に依存することなく最終的にスイッチによって処理されることを保証する、回路切り換えとフレーム切り換えの両方を行なうことのできる、光ファイバースイッチのための転送路割り当てシステムとその方法が必要とされている。

【0009】本発明の目的は、上述した当該分野において周知の従来技術の問題点を解決することである。

【0010】本発明の他の目的は、光ファイバーネットワークの光ファイバースイッチのための高性能な転送路

割り当てシステムとその方法を提供することである。

【0011】本発明の他の目的は、ファイバーチャネル規格にしたがってクラス 1 転送のための回路切り換えとクラス 2 転送のためのフレーム切り換えの両方を効率的に行なうことのできる光ファイバースイッチのためのフレキシブルな転送路割り当てシステムとその方法を提供することである。

【0012】本発明の他の目的は、転送要求がポートが利用できる確率に依存せずスイッチによって最終的に処理されるよう保証することを可能にする光ファイバースイッチのための転送路割り当てシステムとその方法を提供することである。

【0013】

【課題を解決するための手段】つまり、本発明は光ファイバーネットワークにおいて光ファイバースイッチを介した高性能のデータ転送を行なうための新しい転送路割り当てシステムとその方法を提供するものである。この新しい転送路割り当てシステムは次のように構成される。複数のポートがネットワークの複数の光ファイバーチャネルに関係付けられる。それぞれのポートは送信器と受信器からなる対応するポートインテリジェンス機構を有する。

【0014】メモリーインターフェースシステムが複数のポートインテリジェンス機構（あるいはポートに関係付けられ、クラス 2 データ転送（フレーム切り換え）について、出所ポートからの新しい着信データフレームの一時的格納を行なう。このメモリーインターフェースシステムがクラス 1 データ転送（回路切り換え）のためのパ

【0015】スイッチモジュールは、制御分配ネットワーク（CDN）、データ用の主分配ネットワーク（MDN）、およびデータ用の混合分配ネットワーク（IDN）等のいくつかの分配ネットワークからなる。スイッチモジュールはメモリーインターフェースシステム間で選択的にデータ転送路を相互接続することができる。

【0016】セントリは、スイッチモジュールを介してメモリーインターフェースシステムと通信する。セントリはメモリーインターフェースシステムが新しいデータフレームを受け取ったことを判定し、メモリーインターフェースシステムにこの新しいデータフレームの転送路データを要求する。

【0017】スケジューラは、それぞれのポートに対応する宛先待ち行列を維持する。スケジューラはセントリが新しいデータフレームを認識した後セントリから転送路データを受け取る。この転送路データには、たとえば出所ポート標識、メモリアドレス、および宛先ポート標識が含まれる。スケジューラは適当な宛先ポートに対応する宛先待ち行列に転送路データを入れ、各宛先待ち行列からの転送路データをその宛先待ち行列によって定まる順序で検索する。スケジューラは、グルー論理によ

て定義される順序（本実施形態では連続的な順序）で宛先待ち行列をサービスする。

【0018】グルー論理は、ポートインテリジェンス機構と対話して、それぞれのポートに対応する送信器がビジーであるか使用可能（すなわちチャンネルへのデータの送信とスイッチからのデータの受信が可能な状態）であるかを判定する。この問い合わせに基づいて、グルー論理はサービスすべき次の宛先ポートを同定し、スケジューラにそれを知らせる。次に、スケジューラはこの次の宛先ポートに対応する宛先待ち行列にアクセスする。

【0019】アービトラータが、スイッチモジュールを介したデータ転送を最終的に制御し、スケジューラおよびポートインテリジェンス機構と通信する。アービトラータはポートが使用可能であるか、あるいは他のデータ転送要求をサービスするために使用中であるかを判定する。使用可能であれば、アービトラータは主分配ネットワークあるいは混合分配ネットワークを介したポート間でのデータ通信（クラス1転送あるいはクラス2転送）を可能にする。

【0020】本発明の重要な特徴は、このアービトラータが、転送要求が適切な時間にまた最終的に宛先ポートによってサービスされるように保証するための任意の数の公平性保証機構を有する。出所公平性保証機構と宛先公平性保証機構がそれぞれのポートに専用に関係付けられている。それぞれの出所公平性保証機構および公平性保証機構は、対応するポートへの将来のアクセスの順序をデータの出所およびデータの宛先としてそれぞれ確立する。順序の指定には、少なくとも1つのポート標識の指定が含まれる。つまり、それぞれの出所公平性保証機構は将来のアクセスを調整すべき少なくとも1つの宛先ポートを有し、同様に、それぞれの宛先公平性保証機構は将来のアクセスを調整すべき少なくとも1つの出所ポートを有する。

【0021】さらに、本発明の実形態では、複雑性を最小限にし、必要なメモリを少なくするために、それぞれの公平性保証機構はネクスト・イン・ライン標識を用いて構成される。次出所標識は求められた転送要求の直後の、ビジー状態になる転送要求の宛先ポート標識のみを格納する。さらに、次宛先標識は、求められた転送要求の直後の、ビジー状態になる転送要求の出所ポート標識のみを格納する。以上のような方法は、宛先待ち行列がスケジューラによって所定のサービス待ち行列順序（実施形態では連続的な順序）でサービスされることによって可能となるものである。状態の情報に簡単にアクセスすることができる。

【0022】アービトラータにはさらに、ある転送要求がパケット転送を行なうためのものであるが専用転送路（回路切り換え）転送を行なうためのものであるかを判定するためのクラス判定機構を設けることもできる。専用転送路転送の場合には、アービトラータは宛先ポート

がデータの受信と送信の両方が可能な状態であるか、また出所ポートがデータの送出が可能な状態であるかを判定する。

【0023】以上の目的のすべてを達成することに加えて、本発明には多くの利点があり、そのいくつかを次に説明する。

【0024】本発明の利点は、この転送路割り当てシステムは転送路の集中制御を可能にし、その結果ハードウェアが最小限ですみ、転送統計に簡単にアクセスすることができることである。

【0025】本発明の他の利点は、この転送路割り当てシステムと公平性保証機構はともに設計が簡単であり、安価で量産に適し、また信頼性と動作効率が高いことである。

【0026】本発明の他の利点は、この公平性保証機構は光ファイバースwitchの性能とスループットを大幅に向上させることである。

【0027】当業者には、本発明の他の目的、特徴および利点は以下の図面と詳細な説明から明らかになるであろう。かかる目的、特徴および利点も本発明の範囲に含まれるものである。

【0028】

【発明の実施の形態】次に、本発明に係るメモリインタフェースシステムについて、図面を参照して説明する。なお、添付図面において、同一符号は同一部品を指す。図1は、光ファイバースwitch 30の概略図を示す。この光ファイバースwitch 30は、複数の光ファイバチャンネル 32の選択的相互接続を可能にすることによって、光ファイバネットワークの実施を可能にするものである。光ファイバースwitch 30は非常にフレキシブルなシステムであり、ファイバチャンネル規格に基づいたクラス1データ転送のための回路切り換えと、クラス2データ転送のためのフレーム切り換えの両方を可能にし、また他の従来の光ファイバースwitchに比べはるかに高性能なものである。

【0029】アーキテクチャについていえば、光ファイバースwitch 30は複数のチャンネルモジュール 34を有し、このチャンネルモジュール 34に対応するポート（P1～Pi） 33を介して、光ファイバチャンネル 32が接続されている。それぞれのチャンネルモジュール 34は、1つあるいはそれ以上の光ファイバチャンネル 32に接続されている。それぞれのチャンネルモジュール 34は、後に詳しく述べるが、各チャンネルを用いたデータ通信のためのポートインテリジェンス、クラス1データ転送のためのバイパスおよびクラス2データ転送のためのデータフレームの一時的格納のための受信メモリを提供する。チャンネルモジュール 34はスイッチモジュール 36に接続されており、スイッチモジュール 36は電源 37から電気エネルギーを受け取ってこれを分配する。本実施形態では、スイッチモジュール 36

は、バックプレーンの一部として設けられ、多数の機能インターフェース要素を有する。

【0030】スイッチモジュール36は、状態マルチプレクサ41 (MUX) を有する。この状態マルチプレクサ41は、チャンネルモジュール34からポート33とこのポート33に係る回路に関する状態信号を受信するように構成されている。状態信号には少なくとも、そのチャンネルモジュール34に対応する受信メモリ84 (図2) が新しいフレームを受信したことを示す「新フレーム着信」信号と、ポート33から受け取ったデータを受信メモリ84から光ファイバースイッチ30を介して送れるか否かを示す「受信器ready」または「rxready」信号と、混合分配ネットワーク (IDN) 44がデータ転送可能 (使用されていない) かデータ転送不能 (現在使用中) であるかを示す「混合バスready」信号と、ポート33に対応するポートインテリジェンス機構73 (図2) が活動状態であるか不活動状態であるかを示す「ポート活動状態」信号と、ポート33に対応する送信メモリ86 (図2) が光ファイバースイッチ30から (ポート33に宛てられた) データを受信可能であるか否かを示す「送信器ready」信号と、混合分配ネットワーク44が混合転送を実行可能であるか否かを示す「混合ready」信号と、転送路割当システム50にチャンネルモジュール34の対応する状態/制御論理回路85 (図2) から状態情報が転送可能であるか否かを示す「状態転送ready」信号すなわち「xfer ready」信号とが含まれる。

【0031】図1に示すように、主分配ネットワーク (MDN) 42が、光ファイバースイッチ32のデータ転送路を選択的に相互接続する。制御分配ネットワーク (CDN) 43が主分配ネットワーク (MDN) 42を制御し、さまざまなチャンネルモジュール34に制御信号を送る。混合分配ネットワーク (IDN) 44が、チャンネルモジュール34の間の混合転送路を選択的に相互接続する。混合転送路は主分配ネットワーク42に係る付けられたデータ転送路とは別の、主分配ネットワーク42のデータ転送路が使用されているときに選択された光ファイバースイッチ32間のデータの流れを可能にするデータ転送路の集合である。最後に、オプションとしてプロセッサセクタ45を、光ファイバースイッチ30に分散されたプロセッサやコントローラを相互接続するための補助システムの一部として設けることができる。

【0032】転送路割り当てシステム50は、スイッチモジュール36、より詳細には状態マルチプレクサ41、主分配ネットワーク42、制御分配ネットワーク43及び混合分配ネットワーク44に接続されている。転送路割当システム50は、スイッチモジュール36中および光ファイバースイッチ33の間のデータ相互接続路を割り当て、接続の優先順位を決定する。

【0033】また、スイッチモジュール36には、要素コントローラ (EC) 58を接続することもできる。要素コントローラ58は、基本的にはネームサーバ、時間サーバその他のインターフェースシステム30のためのサーバを提供する。要素コントローラ58は、転送路割り当てシステム50との間でサーバ情報の通信を行うためのデータリンク61と、転送路割当システム50との間で状態/制御信号をやりとりするための状態/制御結線62とを有する。また、要素コントローラ58は、結線64を介してチャンネルモジュール34およびマイクロプロセッサセクタ45と、初期化情報や構成情報をやりとりする。

【0034】好ましくは、各チャンネルモジュール34は、図2の概略回路図に示すように構成される。同図に示すように、各チャンネルモジュール34は、メモリインターフェースシステム72と接続されたポートインテリジェンスシステム71を有する。本実施形態では、ポートインテリジェンスシステム71は、1つ或いはそれ以上のポートインテリジェンス機構73を有する。それぞれの光ファイバースイッチ32に、1つのポートインテリジェンス機構73が割り当てられる。それぞれのポートインテリジェンス機構73は、受信器 (RX) 74、送信器 (TX) 76、オブティカルリンクカード (OLC) 75、および状態/制御 (STATCNT) 論理回路85を有する。受信器74と送信器76は、対応する入力光ファイバースイッチ79および出力光ファイバースイッチ83 (図1には、集成的に光ファイバースイッチ32として示す) を介して、ファイバースイッチ規格プロトコルにしたがってチャンネルの固有のビットレートでデータの送受信を行なうようになっている。

【0035】オブティカルリンクカード75は、ポートインテリジェンス機構73を光ファイバースイッチ32に直接インターフェースするのに用いられる。オブティカルリンクカード75は、光ファイバースイッチ32の入力光ファイバースイッチ79と受信器74の間の光・電気変換と、シリアル・パラレル変換を可能にする。さらに、オブティカルリンクカード75は、光ファイバースイッチ32の出力光ファイバースイッチ83と送信器76の間の電気・光変換と、パラレル・シリアル変換を可能にする。オブティカルリンクカード75は、例えば米国IBM社の製造販売するモデルオブティカルリンクカード266や、米国ELDEC社の製造販売するモデルMIM266等の任意の適当なオブティカルリンクカードとすることができる。

【0036】状態/制御論理回路85は、対応する双方向の制御結線87、91によって受信器74と送信器76の両方を監視・制御する。さらに、状態/制御論理回路85は、制御結線95上で制御分配ネットワーク43 (図1) と制御信号をやりとりし、結線96上で状態マルチプレクサ41 (図1) に、例えば対応するポート3

3が利用可能であるか使用中であることを示す状態信号を提供し、結線97を介してメモリインターフェースシステム72に制御信号を送る。さらに、状態/制御論理回路85は、受信器74が新たなフレームを受信するときこれを認識し、その転送クラス1又は転送クラス2、および新しいフレームそれぞれの長さを判定する。たとえば光ファイバスイッチ30(図1)をクラス1データ転送のための双方向の転送路を確保するようにセットアップするために、最初に光ファイバスイッチ30中を送られるSOFc1の場合のように、フレームがデータを持っていないことがある。

【0037】メモリインターフェースシステム72は、ポートインテリジェンスシステム71、特にそこに含まれるそれぞれのポートインテリジェンス機構73に、シリーズ接続あるいはカスケード接続される。メモリインターフェースシステム72は、クラス1バイパスデータ結線98、99によってクラス1データ転送及びクラス2データ転送のため、一時的格納を可能にする。クラス2データ転送のためのデータの格納を行なうために、メモリインターフェースシステム72は、出所データ用の受信メモリ84、宛先データ用の送信メモリ86、および受信メモリ84と送信メモリ86を制御するためのメモリ制御論理回路88を有する。受信メモリ84と送信メモリ86は、必要であれば多数のバッファあるいはメモリブロックに分けることができる。

【0038】メモリインターフェースシステム72がポートインテリジェンスシステム71からクラス1出所データを受け取ると、この出所データは受信メモリ84を迂回して、バイパスデータ結線98、マルチプレクサ66及びデータ結線89に送られる。データ結線89は、出所データをスイッチモジュール36の主分配ネットワーク42または混合分配ネットワーク44のデータバスに送る。メモリ制御論理回路88は、受信器74からクラス1データ転送であるかクラス2データ転送であるかを示すタグ81'を受け取り、それにしたがってクラス制御結線65上でマルチプレクサ66を制御する。受信器74は着信データのヘッダ14(図9)に基づいてタグ81'を生成する。本実施形態では、2ビットのタグが用いられる。タグ「00」は未使用を表わす。タグ「01」はデータを表わす。タグ「10」はクラス1データ転送のSOFあるいはEOFを表わす。タグ「11」はクラス2データ転送のSOFあるいはEOFを表わす。

【0039】メモリ制御論理回路88がタグ81'から判定するように、メモリインターフェースシステム72が着信クラス2出所データ(およびSOFc1フレーム)を受け取ると、受信メモリ84はメモリ制御論理回路88の制御のもとにデータ結線81を介して、受信器74から出所データを読み出し格納する。さらに、タイミングが適当であれば、受信メモリ84は制御論理回路

88の制御のもとに、データ結線67、マルチプレクサ66、およびデータ結線89を介して、データをスイッチモジュール36(図1)の主分配ネットワーク42あるいは混合分配ネットワーク44のデータバスに書き込む。受信メモリ84からデータバスにデータを転送するために、制御分配ネットワーク43は状態/制御論理回路85に制御信号95を送り、状態/制御論理回路85は制御結線97を介して、メモリ制御論理回路88に送出信号を送る。この送出信号は、メモリ制御論理回路88が制御結線92を介して受信メモリ84に送ると共に、メモリ制御回路88がクラス制御結線65を用いてマルチプレクサ66を制御し、このマルチプレクサ66が受信メモリ84から送出信号をデータ結線89に送るようにする。必要であれば、制御分配ネットワーク43は、状態/制御論理回路85に削除信号を送ることによって、受信メモリ84内のフレームを削除することができる。状態/制御論理回路85は、制御結線97を介してメモリ制御論理回路88に削除コマンドを送る。

【0040】主分配ネットワーク42あるいは混合分配ネットワーク44のデータバスから宛先ポート33への宛先データは、データ結線94によって送信メモリ86に送られ、またバイパスデータ結線99によってマルチプレクサ69に送られる。タグ81'上の2ビットタグと同様のタグ94'上の2ビットタグが、この宛先データがクラス1データ転送あるいはクラス2データ転送に対応するとき、これをメモリ制御論理回路88に通知する。クラス1宛先データが受信されると、メモリ制御論理回路88は制御結線68を介して、マルチプレクサ69はデータ結線82を介して、適当なポートインテリジェンス機構73の送信器76にデータを直接送ることによって、送信メモリ86をバイパスするようにマルチプレクサ69を制御する。これに対して、メモリインターフェースシステム72がクラス2宛先データを受け取ると、メモリ制御論理回路88は、この着信宛先データをデータ結線94を介して格納するように、送信メモリ86を制御する。タイミングが適当であれば、この宛先データは最終的にはメモリ制御論理回路88の制御のもとにデータ結線102、マルチプレクサ69、およびデータ結線82を介して適当なポートインテリジェンス機構73の送信器76に送られる。

【0041】図3は、図1のこの新しい転送路割り当てシステム50の一実施形態を詳細に示す。上述したように、転送路割り当てシステム50は、スイッチモジュール36(図1)を介したデータ転送路の割り当てを高性能に実行する。アーキテクチャについていえば、この転送路割り当てシステム50の実施例は、図3に示すように以下の構成要素を有する。すなわち、シーケンサ101が制御結線103を介してCDN43に接続されている。セントリ104が制御結線106を介してCDN43とつながっており、また制御結線107を介してシー

ケンサ 101 に接続されている。タイマ 108 が制御結線 111 を介して セントリ 104 と グルー論理回路 113 に繋がっている。グルー論理 113 は制御結線 112 を介して セントリ 104 と 制御信号をやりとりし、また制御結線 114 を介してシーケンサ 101 と 制御信号をやりとりする。スケジューラ 118 がポート (p1 ~ pi) 33 (あるいはチャンネル 32) のそれぞれに対応する宛先待ち行列 (Qp1 ~ Qpi) 121 を維持する。スケジューラ 118 は結線 117、122 を介して グルー論理 113 と通信する。

【0042】アービトレータ 123 が結線 122 を介してスケジューラ 118 に接続され、結線 116 を介してグルー論理 113 に接続され、結線 126 を介してシーケンサ 101 に接続され、データ結線 124 を介して MDN 42 と IDN 44 と繋がっている。クロザ 128 が結線 129 を介してシーケンサ 101 をモニターし、制御結線 131 を介して CDN 43 から制御信号を受け取り、結線 134 を介して セントリ 104、アービトレータ 123、および要素コントローラ (EC) インターフェース 133 に繋がっている。EC インターフェース 133 はシリアル結線 135 を介して要素コントローラ 58 (図 1) に繋がっている。

【0043】シーケンサ 101 は、任意の適当な論理を用いてたとえば次に説明するような機能を有する従来のフィールドプログラマブルゲートアレイ (FPGA) 内の状態機械として構成することができる。シーケンサ 101 は制御結線 103 を介した CDN 43 への主インターフェースとして機能する。シーケンサ 101 は基本的には セントリ 104、アービトレータ 123 およびクロザ 128 の間で CDN 43 の調整を行なう。

【0044】セントリ 104 は、任意の適当な論理を用いてたとえば FPGA 内の状態機械として構成することができる。セントリ 104 は、状態 MUX 41 (図 1) と状態/制御論理回路 85 (図 2) を介してポート 33 の新フレームの状態 (“新フレーム” 着信信号) をモニターし、スイッチモジュール 36 の MDN 42 を介して経路指定できる新しいフレームがあるかどうかを判定する。基本的には、セントリ 104 は確認されバッファされたフレームのデータを収集し、経路指定テーブルを用いて着信フレームのヘッダ (図 4) から、適当なポート 33 に宛先ポート識別 (DID) をマッピングし、新しいフレームがある宛先ポート 33 を宛先とするものであるか (すなわち、そのフレームをクラス 1 ストリームに混合することができるかどうか) を判定し、また新しいフレームが適正なものであるか誤ったものであるかを判定する。

【0045】セントリ 104 とタイマ 108 は、さらにグルー論理 113 と結線 112、117、122 を介してスケジューラ 118 に待ち行列コマンド情報を与える。待ち行列コマンド情報には、追加信号、フレームビ

ジー (fb sy) 信号、および削除信号 (del) 信号が含まれる。追加信号は、新しいフレームがメモリインターフェースシステム 72 (図 2) の受信メモリ 84 内にあり、光ファイバスイッチ 30 を介して経路指定可能であるとき送出される。セントリ 104 からスケジューラ 118 に追加信号が送られるとき、追加コマンド、宛先ポート (待ち行列)、リンクリスト末尾ポインタ、クラス標識、出所ポート、出所バッファ番号、およびリンクリストヘッダポインタ等の情報が送られる。

10 【0046】fb sy 信号は、受信メモリ 84 に新しいフレームがそのシステムには長すぎると考えられる所定の期間 (すなわち fb sy 期間) 存在したとき送出される。セントリ 104 からスケジューラ 118 に送出される fb sy 信号は、前述したように一般に追加信号と同じ情報を含む。

15 【0047】削除信号は、新しいフレームが受信メモリ 84 内に fb sy 期間より長くフレームの削除を行なうことが可能な他の所定の期間 (すなわち削除期間) 存在したとき送出される。削除信号は他のエラー状態に対しても発することができる。セントリ 104 からスケジューラ 118 に送出されるこの削除信号には、(a) 削除コマンド、(b) 出所ポート、および (c) 出所バッファ番号等の転送路データが含まれる。

20 【0048】タイマ 108 は、デジタル信号プロセッサ (DSP) 等の従来の任意の処理機構を用いて構成することができる。タイマ 108 はある新しいフレームが、あるチャンネルモジュール 32 の受信メモリ 84 内に存在する時間を測定し、セントリ 104 から fb sy 信号および削除信号をいつ発するかを決定する。このため、
30 タイマ 108 は内部にそれぞれの新しいフレームの fb sy 期間と削除期間を追跡するための fb sy/削除クロックを保持する。タイマ 108 は、新しいフレームをポート 33 から転送することができるとき結線 111 上でセントリ 104 から fb sy/削除クロックを起動するための開始 (INIT) 信号 111 を受け取り、fb sy/削除クロックをクリアするためにグルー論理 113 と結線 116、111 を介してアービトレータ 123 からタイマクリア (clr) 信号を受け取り、アービトレータ 123 がタイマクリア信号を受け取っていない場合、fb sy 期間と削除期間の経過後に結線 111 を介してセントリ 104 に fb sy 信号と削除信号を出力する。

40 【0049】グルー論理 113 は、主としてセントリ 104、タイマ 108、シーケンサ 101、アービトレータ 123、およびスケジューラ 118 の間のインターフェースとして機能する。本実施形態では、グルー論理 113 は FPGA あるいは他の適当な論理を用いて状態機械として実施される。また、グルー論理 113 は本実施例では他の機能を実行する。グルー論理 113 はシーケンサ 101 を介してポートインテリジェンス機構 73 が
50

データフレームを受け取れる状態であることを示す送信 ready (txready) 信号と混合 ready 信号をモニターする。グルー論理 113 は循環探索シーケンスを実行する。このシーケンスでは発せられない tx ready 信号および混合 ready 信号をマスクすることによって発せられた tx ready 信号および混合 ready 信号が順次探索され、これによってデータを受け取るべき次の宛先ポート 33 (サービスすべき次の宛先待ち行列 122 のように) が判定される。グルー論理 113 は結線 117、122 を介してスケジューラ 118 にサービスすべき次の宛先ポートの識別を送る。このとき、グルー論理 113 はスケジューラ 118 に二重リンクリスト待ち行列の先頭を同定するベースポインタアドレスと特定の宛先待ち行列を同定する宛先待ち行列標識からなるワードを送る。

【0050】DSP等の任意の従来の処理機構を用いて構成することのできるスケジューラ 118 は、宛先待ち行列 (Qp1~Qpi) 121 を維持・管理し、セントリ 104 から待ち行列管理コマンド、詳細には追加信号、fbsy 信号、および削除信号を受け取る。また、スケジューラ 118 は、グルー論理 113 から結線 117、122 を介して次宛先ポート信号を受け取る。この次宛先ポート信号はサービスすべき次のポート 33 (したがって、サービスすべき次の宛先待ち行列 121) を示す。

【0051】スケジューラ 118 はそれぞれの光ファイバーポート (p1~pi) 33 に対する宛先待ち行列 (Qp1~Qpi) 121 を維持・更新し、対応する各ポート 33 に関係する待ち行列エントリを格納するように構成される。それぞれの待ち行列エントリは、(a) データの転送元となる出所ポート 33、(b) そのフレームが位置する受信メモリ 84 内の出所バッファ番号、および (c) その待ち行列エントリが通常のフレームエントリ、クラス 1 データ転送の始点であるフレーム ("SOFc1") のエントリ、待ち行列先頭エントリ、待ち行列末尾エントリのいずれであるかを示すフラグを同定する転送路データを有する。それぞれの待ち行列は、リンクリスト、好適には二重あるいは複数リンクリストによって関係付けられ、それぞれの待ち行列エントリは、(前の出所ポートとバッファ番号を用いて) 前の待ち行列エントリを指定する後方リンクリストおよび (次の出所ポートとバッファ番号を用いて) 後続の待ち行列エントリを指定する前方リンクと関係付けられる。二重リンクリストは最小限のハードウェアで待ち行列を構築し操作する簡単な方法を提供するものである。

【0052】スケジューラ 118 が、セントリ 104 からの追加信号によってポートインテリジェンス機構 73 からの新しいフレームに関する転送路データを受け取ると、スケジューラ 118 はこの転送路データを前方および後方リンクとともに適当な宛先待ち行列 121 に格納

する。

【0053】スケジューラ 118 は、グルー論理 113 によって定められる順序にしたがって待ち行列 121 からのデータを書き出す。さらに、スケジューラ 118 が特定の宛先待ち行列 121 から転送路データを書き出すとき、スケジューラ 118 はこの宛先待ち行列 121 から対応するリンクリストによって定められる順序にしたがって待ち行列エントリを検索し、出所ポート 33 と宛先ポート 33 への新しいフレームの転送の要求をアービトレータ 123 に送る。アービトレータ 123 は、関係するポートが使用可能であればスイッチモジュール 36 に MDN 42 を介して適当なデータ転送路を相互接続させる。

【0054】スケジューラ 118 が、セントリ 104 からのあるフレームに対応する fbsy 信号を受け取ると、スケジューラ 118 はこの fbsy 信号に対応する待ち行列エントリを取り、この待ち行列エントリをそのデータの出所であるポート 33 に対応する宛先待ち行列 121 に移動して、そのデータフレームが最終的にもとのポート 33 に戻されるようにする。

【0055】スケジューラ 118 が、あるフレームに対応する削除信号を受け取ると、スケジューラ 118 はそのフレームに対応する待ち行列エントリを削除する。この削除されるフレームは前にセントリ 104 からの fbsy 信号によってもとの出所ポート 33 に対応する待ち行列に入れられたものであることに注意しなければならない。

【0056】アービトレータ 123 は、任意の適当な論理、好適には FPG A を用いて構成される状態機械で構成される。アービトレータ 123 はポートインテリジェンス機構 73 の状態を追跡し、ポートインテリジェンス機構 73 がデータの送受信を行なえるかどうかを判定し、ポートインテリジェンス機構 73 間の接続を調整する。すなわち、アービトレータ 123 はポートインテリジェンス機構 73 の状態/制御論理 85 (図 2) によって生成される送信 ready (txready) 信号、混合 ready 信号、混合パス ready 信号、および受信 ready (rxready) 信号をモニターする。また、アービトレータ 123 が特定の出所ポート 33 から特定の宛先ポート 33 へデータを転送するための転送要求信号をスケジューラ 118 から受け取ると、アービトレータ 123 はこの転送がクラス 1 (回路切り換え) であるかクラス 2 (フレーム切り換え) であるかを判定する。

【0057】クラス 1 である場合、アービトレータ 123 はこの特定の出所ポート 33 と宛先ポート 33 がデータの送信と受信の両方に使用できるかどうかを判定する。これはクラス 1 転送は双方向であり、スイッチ 30 を介した専用のデータ転送路を必要とするためである。クラス 2 である場合、アービトレータ 123 はこの特定

の出所ポート 33 が転送可能であるか、また特定の宛先ポート 33 がデータフレームを受信できる状態であるかどうかを判定する。

【0058】ポート 33 が転送（クラス 1 あるいは 2）を行なえる状態ではないとき、アービトラータ 123 はスケジューラの要求を拒絶し、スケジューラ 118 は後で転送要求が最終的に認められるまで任意の適当なアルゴリズムを用いて再度この転送要求を送る。アービトラータ 123 によって転送要求が認められると、アービトラータ 123 は MDN 42 が適当な出所ポート 33 を適当な宛先ポート 33 に接続するように MDN 42（図 1）を制御する。さらに、データ転送要求がアービトラータによって認められたとき、アービトラータ 123 はスケジューラ 118 にスケジューラクリア信号を送り、スケジューラ 118 に関するリンクリスト待ち行列から関係する待ち行列エントリを削除させる。また、アービトラータ 123 はタイマ 108 にタイマクリア信号を送って転送されるデータフレームに対するタイマの内部 busy/削除クロックをクリアする。

【0059】アービトラータ 108 は、さまざまな態様で構成して上述した機能を達成することができる。しかし、アービトラータ 108 の好適なアーキテクチャと方法論は次に述べるとおりである。ここに説明するアービトラータ 108 は、スケジューラ 104 からの要求 121 が 1 つも拒絶されることのないように、資源の割り当ての公平性を保証するための公平性保証アルゴリズムを実施する点で好適である。つまり、このアービトラータ 108 は、すべての転送要求が適切な時間にサービスされ、確率に依存しないようにするものである。

【0060】次に、公平性保証アルゴリズムについて説明する。この公平性保証アルゴリズムは、光ファイバスイッチ 30 を介したあらゆるデータ転送要求が常にサービスされるようにし、それによって対応するデータがスイッチ 30 を介して 1 つのポート 33 から他のポート 33 に経路指定されるようにする。従来の多くの場合のように、各ポート 33 が厳密に利用可能性に応じて割り当てられる場合、所望のポート 33 へのアクセスが常に拒絶されるデータが出てくる。この問題はポートアクセスのプロトコルにおける共鳴あるいは繰り返しパターンによってより深刻になる場合がある。すなわち、長いシーケンスによって 1 つあるいはそれ以上のポート 33 への非常に規則的で周期的なトラヒックパターンが生じる場合がある。このパターンがポート資源の割り当てに一致する場合、このシーケンスはポート資源の調整を不当に終了し、他のシーケンスを遅くする可能性がある。

【0061】アービトラータ 123 内に採用することのできる非常に有効な公平性保証アルゴリズムの一例を説明する。この公平性保証アルゴリズムは、拒絶された転送要求を含むポートへのアクセスのためのシーケンスを確立することによってデータが最終的にあるポートに転

送されることを保証するものである。このシーケンスはたとえば待ち行列あるいはその他の表示機構を用いて実施することができる。さらに、このシーケンスは任意の所望の順序付けアルゴリズムあるいは順序付け法を実行することができる。

【0062】しかし、本実施形態では、このシーケンスは 1 つの“ネクスト・イン・ライン（次）” 標識を用いて簡単に、また有効に実施することができる。すなわち、“次” 宛先標識が宛先ポートとしてのある特定のポートについて“次の” 転送要求を同定する。また、出所ポートとしてのそれぞれのポートについて、“次の” 転送要求を同定する“次” 出所標識がある。基本的には、あるフレーム転送要求があるポート（出所あるいは宛先ポート）へのアクセスを拒絶され、そのポートへのアクセスを待っている他のフレーム転送要求がない場合、この拒絶されたフレーム転送要求は、そのポートにアクセスすべき“次の” 転送要求となる。同じポートを必要とする後続のフレーム転送要求は、そのポートがこれら後続のフレーム転送要求のあった時点で使用可能であるか否かにかかわらずアクセスを拒絶される。このシステムは次の次転送要求のみを格納し、後続のフレーム転送要求を格納しないことによって機能する。これは、宛先待ち行列は所定の順序で（本実施例では順に）サービスされるためである。最後に、次転送要求の識別のみが保持されるため、必要なメモリと複雑性が最小限になることに注意しなければならない。

【0063】後に詳述するように、クラス 1 およびクラス 2 の両方のデータ転送を可能にするためには、ポート 33 に関係付けられるネクスト・イン・ライン標識（宛先および出所）は次のように設定される。あるポート 33 の次宛先標識は（a）そのポートの宛先待ち行列内のあるフレームが次の宛先であることを示す“dstQ” あるいは（b）そのポートの rx（ピッチャ）側のクラス 1 転送を示す“SOFc1” のいずれかに設定することができる。さらに、あるポート 33 の次出所標識は（a）“空” あるいは（b）特定の宛先待ち行列 121 のいずれかに設定することができる。

【0064】図 4 にはこの公平性保証アルゴリズム、より詳細には参照符号 130 で示すネクスト・イン・ライン公平性保証アルゴリズムを理解するためのデータ経路指定法の一例を示す。同図において、矢印 130 で示すデータ転送が現在進行中であり、同時に他の 4 つのデータフレームが、さまざまなポート 33 へのアクセスをめぐって光ファイバスイッチ 30 内で競合している。スケジューラ待ち行列 121 に示すように、ポート p1 からポート p2 へのあるフレーム（待ち行列 Qp2 内のフレーム“1, 1”；最初の桁は出所ポートを示し、第 2 の桁はバッファすなわち出所ポートにおけるメモリブロックを示す）がポート p1 からポート p3 への他のフレーム（待ち行列 Qp3 内のフレーム“1, 2”）と競合す

る。しかし、ポート p 1 はすでにフレームを転送中であるため、上記の転送要求はいずれもこの時点では認められない。待ち行列サービスは使用可能な宛先ポート 3 3 に対応する待ち行列に順次提供されるため、待ち行列 Qp 3 は転送許可が与えられた待ち行列 Qp 2 の直後にサービスされる。これを参照符号 1 3 1 で示す。したがって、待ち行列 Qp 3 は出所ポートとしてのポート p 1 については次の出所であり、p 1 の次出所ポート標識は待ち行列 Qp 3 を示す（簡単に示せば“次の SP [1] = 3”）。

【0065】クラス 1 データ転送（図 4 には“SOF c 1”で表わす）の場合、状況はより複雑である。クラス 1 データ転送には、宛先ポート 3 3 の tx（データを受け取る）側だけでなく、その出所ポート 3 3 の tx 側をも必要とする。したがって、出所ポート p 3 からポート p 1 へのクラス 1 データ転送は、ポート p 3 を宛先とするフレームが転送されるまで待たねばならない。これは、ポート p 3 の次宛先標識を（そのポートの rx（ピッチャ）側のクラス 1 データ転送を示す）“SOF c 1”にセットすることによって、宛先ポートとしてのポート p 3 の次の宛先としてポート p 3 において SOF c 1 転送を確立することによって行なわれる。さらに、図 4 の待ち行列 Qp 3 内のフレーム“1, 2”もまた宛先ポートとしてのポート p 3 の次の宛先であり、これは“次の DP [3] = dstQ”で示される。

【0066】公平性保証アルゴリズムが採用されると、スケジューラ 1 1 8 内の待ち行列 1 2 1 からの待ち行列エントリの削除は特に煩雑になる。たとえば、待ち行列 Qp 3 内のフレーム“1, 2”が削除される場合を考える。この場合、フレーム“1, 2”は待ち行列 Qp 3 の先頭から削除するだけでなく、これらのポートに対応する次出所標識および宛先標識からも削除しなければならない。これを行なわないと、“デッドリ・エンブラス”が起ることがある。たとえば、待ち行列 Qp 1 内のフレーム“3, 1”（すなわち SOF c 1 転送）をポート p 3 の次出所標識（すなわち次の SP [3] = 1）で指定したとする。しかし、これは次の DP [3] = dstQ によって阻止される。待ち行列 Qp 3 のフレーム“1, 2”が削除されると、ポート p 3 を宛先とするフレーム“3, 2”は次の宛先にアクセスすることができるが、その所望の出所ポートへのアクセスを得ることはできない。さらに、次の出所ポートへのアクセスを持つフレーム“3, 1”は所望の次宛先ポートを得ることができない。この状態は無期限に続く場合がある。

【0067】次に、公平性保証機構を用いたアービトラータについて説明する。図 5 から図 8 は、その動作のさまざまな段階で前述した公平性保証アルゴリズムを採用したアービトラータ 1 0 8 の、アーキテクチャと方法を示すフローチャートである。これらの図に示す機能は、ハードウェア、ソフトウェア、あるいはその両者の組み合わせによって達成することができる。しかし、本実施

形態では、アービトラータ 1 0 8 は多数のレジスタ/メモリアレーとつながった多数の論理状態機械として高速ハードウェア内に構成される。すなわち、状態機械を用いて図 5 から図 8 のそれぞれのフローチャートに示す方法が実行され、レジスタ/メモリアレーは、次出所標識および宛先標識および次に説明するその他の必要なデータを格納するように構成される。

【0068】図 5 のフローチャート 1 3 2 は、高レベルの方法論を示し、結線 1 2 2（図 3）を介してスケジューラ 1 0 4 から発せられるそれぞれの転送要求についてアービトラータ 1 2 3 によって実行される。図 5 から図 8 に示すフローチャートでは、アービトラータ 1 2 3 は転送要求に対して MDN 4 2（“MDN 許可”）あるいは IDN 4 4（“IDN 許可”）の使用を認めることができる。アービトラータ 1 2 3 は転送要求に対して、宛先としてのポート 3 3 の使用（“dpNotFree”）、出所としてのポート 3 3 の使用（“spNotFree”）、および IDN 4 4 の使用（“imixNotFree”）を拒否することがある。

【0069】まず、フローチャートブロック 1 3 3 に示すように、スケジューラ 1 1 8 からの転送要求がデータの宛先または出所として要素コントローラ（EC）5 8 を含むかどうか判定される（“dp=EC 又は sp=EC？”）。含まれる場合、フローチャートブロック 1 3 4 に示すように、図 6 に示す方法がアービトラータ 1 0 8 によって実行される。

【0070】データ転送要求が EC 5 8 に関係しない場合、アービトラータ 1 0 8 はフローチャートブロック 1 3 5 に示すように転送要求が IDN 4 4 を使用するものであるかを判定する。基本的には、問題は出所ポートあるいは宛先ポートがデータ転送路の確保を必要とするクラス 1 転送にすでに関係しているかどうかである（“クラス 1 [dp] 又は クラス [sp] 1？”）。IDN 4 4 は MDN 4 2 がクラス 1 転送のために使用できないときデータの転送に用いられる。混合接続が発生する場合、アービトラータはフローチャートブロック 1 3 6 に示すように、図 7 に示す方法を実行する。

【0071】IDN 4 4 がこの転送要求に使用されない場合、アービトラータ 1 0 8 はフローチャートブロック 1 3 7 に示すように MDN 4 2 上にクラス 1 接続を確立すべきかどうかを判定する（“SOF c 1？”）。クラス 1 接続を行なわなければならない場合、アービトラータはフローチャート 1 3 8 に示すように図 8 に示す方法を実行する。

【0072】次に、アービトラータ 1 0 8 は宛先としてのこの宛先ポートについて宛先公平性保証機構 1 3 0' を採用する。フローチャートブロック 1 3 9（“次の DP [dp] = SOF c 1？”）に示すように、この転送要求の宛先ポートに対応する次宛先標識がクラス 1 接続を示すかどうか判定される。YES であれば、この転送要求

はフローチャートブロック 1 4 1 に示すように拒絶される (“dpNotFree”)。これは、前の要求が優先権を持ったためである。

【0073】この次宛先標識は、宛先ポートを指定していない場合、フローチャートブロック 1 4 2 に示すように、アービトラータ 1 0 8 はこの次宛先標識を、サービスされている現在の待ち行列を指定する、すなわちそのポートの待ち行列内のあるフレームが次の宛先であることを指定するように設定する。

【0074】次に、アービトラータ 1 0 8 は出所としてのこの出所ポートに対して出所公平性保証機構 1 3 0”を採用する。前の要求が同じ出所ポートを出所とするものである場合、この前の要求は現在の要求の前にアクセスを認められる。そうでない場合、現在の要求は出所としてこの特定の出所ポートへのアクセスを認められる。より詳細には、フローチャートブロック 1 4 3 (“次の SP[sp] != NULL & 次の SP[sp] != dp?”) において、現在の転送要求によって指定された出所ポートのネクスト・イン・ライン標識が空であるか、あるいは次ポートがすでに指定されているかが判定される。すでに指定されている場合、この出所ポートへのアクセスはフローチャートブロック 1 4 4 に示すように拒絶される (“spNotFree”)。空である場合、この出所ポートの次出所標識はこの転送要求に指定され、フチャートブロック 1 4 5 に示すように現在サービスされている宛先ポートに対応するように設定される (“次の SP [sp] = dp”)。

【0075】次に、フローチャートブロック 1 4 6 に示すように現在の出所ポートがデータの送出が可能であるかどうか判定される (“sPFree [sp] ?”)。基本的には、この特定の出所ポートからの rxready 信号が調べられる。送出可能でない場合、フローチャートブロック 1 4 4 に示すように、この出所ポートは使用不能であり、転送要求は拒絶される (“spNotFree”)。使用可能である場合、アービトラータ 1 0 8 はレートマッチングを行なう。本発明では、レートマッチング機能はオプションであり、光ファイバースイッチ 3 0 (図 1) のスループットレートを操作する能力を必要とする。光ファイバースイッチ 3 0 (図 1) を介してデータが搬送される速度は MDN 4 2 内に一連のバラレルデータバスを設け、データを同時に転送する MDN 4 2 のバスの数を設定することによって操作することができる。

【0076】このレートマッチング法は、フローチャートブロック 1 4 7 から 1 4 9 に示すように実行される。出所ポートのビットレートが宛先ポートのビットレートより大きい場合、MDN 4 2 を介した転送のビットレートが宛先ポートのレートとして設定される。あるいは、このレートは出所ポートのレートに設定される。最後に、フローチャートブロック 1 5 0 に示すように MDN 4 2 へのアクセスが認められる (“MDN許可”)。

【0077】図 6 は、EC 5 8 へのアクセス要求を含む

フローチャートブロック 1 3 4 に示す方法を説明するものである。EC 5 8 へのアクセスは IDN 4 4 を介して発生する。まず、図 6 に示すように、EC 5 8 がフローチャートブロック 1 5 1 に示すように所期の出所ポートであるか宛先ポートであるかが判定される (“EC=dp?”)。EC 5 8 が所期の宛先ポートである場合、公平性を保証するために出所としてのこの出所ポートに対して出所公平性保証機構 1 3 0” が用いられる。前の要求が同じ出所ポートを宛先とするものである場合、この前の要求は現在の要求の前にアクセスを認められる。そうでない場合、現在の要求は出所としてこの特定の出所へのアクセスを認められる。

【0078】このとき、フローチャートブロック 1 5 2 に示すように、EC (出所ポート) の次出所標識が空であるかあるいはすでに指定されているかどうか判定される (“次の SP[sp] != sp=NULL & 次の SP[sp] != dp?”)。すでに指定されている場合、フローチャートブロック 1 5 3 に示すようにこの要求は拒絶される (“ImixNotFree”)。空である場合、EC 5 8 の次出所標識はフローチャートブロック 1 5 4 に示すように現在の要求に設定される (“次の SP[sp]=dp”)。

【0079】次に、IDN 4 4 が使用可能であるか (“ImixRX[sp]?”) またこの出所ポートが使用可能であるか (“sPFree[sp]?”) が照会される。使用可能でなければ、フローチャートブロック 1 5 3 に示すように IDN 4 4 へのアクセスは拒絶される (“ImixNotFree”)。使用可能である場合、フローチャートブロック 1 5 6 に示すように、この要求がクラス 1 データ転送を含むものであるかどうか判定される (“SOFc 1”)。クラス 1 データ転送を行なう必要がない場合、この転送要求は IDN へのアクセスを認められ (“IDN許可”)、クラス 2 データ転送が発生する。クラス 1 転送を行なわなければならない場合、アービトラータ 1 0 8 によってフラグがセットされ、このフラグがスイッチモジュール 3 6 に送られ、スイッチモジュール 3 6 が双方向のクラス 1 データフローを可能にするようにセットアップされ、IDN 4 4 へのアクセスが認められる (“IDN許可”)。

【0080】フローチャートブロック 1 5 1 に戻ると、EC 5 8 が出所ポートである場合、フローチャートブロック 1 5 8 に示すように、この出所ポートが IDN を使用するものであるか、すでにクラス 1 転送に関与しているかどうか判定される (“dPImixが SOFc 1 And クラス 1 [dp]?”)。上記のいずれかが肯定されると、フローチャートブロック 1 5 9 に示すようにこの転送要求は拒絶される (“dpNotFree”)。否定されると、処理はフローチャートブロック 1 6 1 に移行する。

【0081】フローチャートブロック 1 6 1 から、アービトラータ 1 2 3 は宛先としての EC 5 8 (すなわち宛先ポート) に対して宛先公平性保証機構 1 3 0’ を用い

る。ブロック 161 に示すように、この宛先ポートに対応する次宛先標識が前の転送要求からのクラス 1 データ転送を指定するものであるかどうかを照会される。YES である場合、フローチャートブロック 159 に示すようにこの要求は拒絶される (“spNotFree”)。NO である場合、EC 58 に対応する次宛先標識はフローチャートブロック 162 に示すように現在の転送要求の宛先待ち行列に設定される (“次の DP[dp]=dstQ”)。

【0082】さらに、フローチャートブロック 163 に示すように、IDN 44 がこの宛先ポートへのデータ転送に使用できるかどうか判定される (“imixTc[dp]?”)。基本的には、IDN 44 からのこの宛先ポートに対応する txready 信号が調べられる。使用可能でない場合、フローチャートブロック 164 に示すように、アクセスは拒絶される (“ImixNotFree”)。使用可能である場合、フローチャートブロック 165 に示すように、クラス 1 転送を確立すべきであるかどうか判定される (“SOFc1?”)。クラス 1 接続を確立すべき場合、フローチャートブロック 166 に示すように、スイッチモジュール 36 を構成するようにフラグがセットされ (“クラス 1[dp]=1”)、IDN 44 へのアクセスが認められる (“IDN 許可”)。クラス 1 接続を確立する必要がない場合、IDN 44 を介してクラス 2 接続が発生する。

【0083】IDN 44 にアクセスする方法は図 7 に示し、次に説明する通りである。まず、フローチャートブロック 168 に示すように、この転送要求がクラス 1 転送あるいはクラス 2 転送のどちらを要求するものであるかが判定される (“SOFc1?”)。クラス 1 転送を行なうべき場合、フローチャートブロック 169 に示すように、出所ポートがすでにクラス 1 転送状態であるかどうか判定される (“クラス [sp]?”)。YES である場合、フローチャートブロック 171 に示すようにこの要求は誤りであり、拒絶される (“ディスク SC I”)。フローチャートブロック 169 でこの出所ポートがクラス 1 接続されていないと判定されると、アクセスは拒絶される。これは、フローチャートブロック 172 に示すように、この宛先ポートは空いていないためである (“spNotFree”)。

【0084】フローチャートブロック 168 においてクラス 2 接続を確立すべきことが判定される場合、次にフローチャートブロック 173 に示すように、この宛先ポートが混合転送、クラス 1 転送のいずれに関与しているかが照会される (“!dpImix& クラス 1[dp]?”)。YES である場合、フローチャートブロック 172 に示すようにこの宛先ポートは使用不能であることからこの転送要求は拒絶される (“dpNotFree”)。

【0085】NO である場合、アービトラータ 123 は宛先公平性保証機構 130' を用いてこの宛先ポートの宛先としての公平性を保証する。この宛先ポートを宛先

として使用する要求がそれ以前に存在した場合、この前の要求が現在の要求に先だって認められる。この場合、フローチャートブロック 174 に示すように、この転送要求の宛先ポートに対応する次宛先標識がクラス 1 転送を示しているかどうかを照会される (“次の DP[dp]=SOFc1”)。YES である場合、この宛先ポートは空いておらず (“dpNotFree”)、この転送要求は公平性保証機構 y 前記通常モードされる。NO である場合、処理は続行され、フローチャートブロック 175 に示すようにこの宛先ポートの次宛先標識はこの転送要求に示される現在の待ち行列に設定される (“次の DP[dp]=dstQ” Z)。

【0086】出所公平性保証アルゴリズム 130' がこの出所ポートの出所としての公平性を判定する。前の要求が同じ出所ポートを出所とするものであった場合、この前の要求が現在の要求に先だってアクセスを認められる。そうでない場合、現在の要求がこの特定の出所ポートを出所とするアクセスを求められる。より詳細には、フローチャートブロック 176 に示すように (“次の SP[sp]!=NULL & 次の SP[sp]!=dp?”)、この転送要求によって同定される出所ポートの次出所標識が空であるか指定されているかが判定される。すでに指定されている場合、フローチャートブロック 177 に示すように、IDN 44 は使用不能であり、この転送要求は拒絶される (“ImixNotFree”)。空である場合、次出所標識はフローチャートブロック 177 に示すように現在の転送要求によって同定される宛先ポートに設定される (“次の SP[sp]=dp”)。

【0087】最後に、フローチャートブロック 178 に示すように、IDN 44 がこの特定の出所ポートから特定の宛先ポートにデータの転送を行なえる状態であるかどうか、また出所が現在空いているかどうか判定される (“imixRX[sp] & imixTX[dp] & sp Free[sp]”)。このとき、出所ポートに対応する IDN 44 の rxready と宛先ポートに対応する IDN 44 の txready が調べられる。YES である場合、この要求は IDN 44 へのアクセスを認められる (“IDN 許可”)。NO である場合、フローチャートブロック 177 に示すようにこの転送要求は拒絶される (“ImixNotFree”)。

【0088】クラス 1 接続に対する要求を実行する方法は図 8 のフローチャートに示し、次に説明する通りである。まず、フローチャートブロック 181 (“次の DP[dp]=SOFc1”)において、アービトラータ 123 が宛先公平性保証機構 130' を用いてこの宛先ポートの宛先としての公平性を保証する。すなわち、この宛先ポートの次宛先標識がすでにクラス 1 接続に設定されているかどうか判定される。YES である場合、フローチャートブロック 182 に示すように、この特定の宛先ポートへのアクセスと転送要求は拒絶される (“dpNotFree

e”）。宛先ポート標識からこの宛先ポートが現在クラス1接続されていないと判定される場合、この宛先ポートの次宛先標識はフローチャートブロック183に示すようにこの転送要求に対応する現在の待ち行列に設定される(“次のDP[dp]=dstQ”)。

【0089】次に、アービトラータ123は宛先公平性保証機構130'を用いてこの出所ポートの宛先としての公平性を保証する。この場合、フローチャートブロック184に示すように、現在の転送要求によって同定される出所ポートの次宛先標識が、現在の転送要求を出した現在の宛先待ち行列に設定されているかどうか照会される(“次のDP[sp]=dstQ?”)。YESである場合、この出所ポートは空いておらず、フローチャートブロック185に示すようにこの転送要求は拒絶される(“spNotFree”)。NOである場合、フローチャートブロック186に示すように、現在の出所ポートの次宛先標識はクラス1転送として指定される。

【0090】さらに、出所公平性保証機構130”を用いてこの宛先ポートの出所としての公平性が保証される。フローチャートブロック187に示すように、この宛先ポートの次出所標識が空であるか、すでにある宛先ポートに指定されているかが判定される(“次のSP[dp]!=Null & 次のSP[dp]!=DP?”)。YESである場合、ブロック182でこの転送要求を拒絶することによってこの出所ポートへのアクセスが拒絶される(“dpNotFree”)。まだ指定されていない場合、フローチャートブロック188に示すように、この宛先ポートの次出所標識がこの転送要求の現在の宛先ポートに設定される(“次のSP[dp]=dp”)。

【0091】次に、この出所ポートについて出所公平性保証機構130”が用いられる。このアルゴリズムはフローチャートブロック189から始まり、このブロックでは、この出所ポートの次出所標識が空であるか、使用されているかが判定される。使用されている場合、フローチャートブロック185に示すようにこの出所ポートは空いておらず、転送要求は拒絶される(“spNotFree”)。空である場合、この出所ポートの次出所標識はブロック191に示すようにこの転送要求の現在の宛先ポートに設定される(“次のSP[sp]=dp”)。

【0092】フローチャートブロック192に示すように、この出所ポートが宛先ポートとして使用可能であるか(“spFree[sp]?”)、また宛先ポートは出所ポートとして使用可能であるか(“spFree[dp]?”)が判定される。NOである場合、フローチャートブロック185に示すようにこの転送要求は拒絶される(“spNotFree”)。YESである場合、フローチャートブロック193に示すように、この出所ポートが宛先ポートとして使用可能であるかどうか判定される(“dpFree[sp]?”)。NOである場合、フローチャートブロック182に示すように、この転送要求は拒絶される(“dpNotF

ree”)。そうでない場合、この転送要求が認められ(“MDN許可”)、MDN42へのアクセスが発生する。

【0093】クローザ128は、任意の適当な論理、好適には市販のFPGAを用いた状態機械で構成される。クローザ128はCDN43への結線131を介してポート33に関する送信状態ready信号(xref ready: 状態情報の転送が可能であるか否かを示す)をモニターする。クローザ128はシーケンサ101にインターフェースして、CDN切り換えコマンドを実行し、また状態信号の集合について宛先ポートコマンドを実行する。さらに、クローザ128はフレーム統計をとり、ECインターフェース133に増分カウントコマンドを送る。この統計と増分カウントは高レベル最適化アルゴリズムによるネットワークの性能とネットワーク管理の最適化に用いられる。

【0094】ECインターフェース133はDSP等の任意の従来の処理機構を用いて構成される。ECインターフェース133はシリアル結線135を介して転送路割り当てシステム50を要素コントローラ58(図1)にインターフェースする。ECインターフェース133は要素コントローラ58からのコマンドを受け取り、このコマンドからバケットを構築し、要素コントローラ58からの読み出しあるいは書き込みコマンドを実行するように転送路割り当てシステムのメモリの維持とアクセスを行ない、アドレス確認テーブル情報を変更するためにセントリ104にコマンドを送り、専用のシリアルポート135を介してタイマ108にコマンドを送る。

【0095】次に、上述の転送路割り当てシステム50の動作を説明する。新しい転送路割り当てシステム50を有する光ファイバスイッチ30の全体的動作を、図1から図3を参照して説明する。より詳細には、新しいデータフレームが光ファイバスイッチ30を介して経路指定されるクラス2データ転送とポート33の間で双方向の専用のデータ転送路が確立されるクラス2データ転送を参照して説明する。

【0096】まず、データフレームは出所ポート(p1~pi)33からチャンネルモジュール32の受信メモリ84に送られる。この出所ポート33に対応する状態/制御論理85が状態MUX41にrxready信号を出力し、状態MUX41はこのrxready信号をセントリ104に送る。このrxready信号に基づいて、セントリ104は新しいフレームがあることを認識し、シーケンサ101にこの新しいフレームの出所ポート33に対応する状態/制御論理85に対して転送路データを要求するように要求する。この転送路データには、出所識別(SID)、受信メモリ84におけるこのフレームの場所を表わすバッファ番号、宛先ポート識別(DID)、およびクラス標識(クラス1あるいは2)が含まれる。

【0097】シーケンサ101はCDN43に一連の制御コマンドを送って、CDNをその出所ポート33に対応する状態/制御論理85にこの新しいフレームに関係付けられた転送路データを送らせるようにセットアップする。すると、CDN43はこの新しいフレームの転送路データを状態/制御論理85からセントリ104に転送させる。セントリ104はSIDを確認し、DIDを特定の宛先ポート33にマッピングする。DIDかSIDのいずれかが無効である場合、そのフレーム自体の代わりにフレーム拒絶信号が状態/制御論理85の待ち行列に入れられる。

【0098】セントリ104は、タイマ108とスケジューラ118の両方に新しいフレームがあることを知らせる。タイマはbsy/削除クロックを起動する。さらに、セントリ104は同時にスケジューラ118に追加信号を送る。スケジューラ118はセントリ104からの追加信号に関係付けられた宛先待ち行列標識に基づいて使用すべき待ち行列を判定する。この追加信号中の末尾ベースポインタアドレスは、この新しいフレームに対応する転送路データが追加される所望の宛先待ち行列の最後尾にアクセスするのに用いられる。スケジューラ118は出所ポート標識、(受信メモリ84内の)出所バッファ番号、およびクラス標識を用いて特定の宛先待ち行列内に待ち行列エントリを生成し、この待ち行列エントリはこの特定の宛先待ち行列121の末尾に入れられる。

【0099】タイマ108が、アービトラータ123からタイマクリア信号を受け取る前にタイマ108においてfbsy期間が経過すると、タイマ108はセントリ104にfbsy信号を送り、セントリ104はスケジューラ118にfbsy信号116を送る。このとき、スケジューラ118はこのfbsy信号に対応する待ち行列エントリを取り、この待ち行列エントリをデータの出所であるポート33に対応する待ち行列に移動し、これによってこのデータは最終的にはその出所ポートに返される。

【0100】タイマ108が、アービトラータ123からタイマクリア信号を受け取る前に、タイマ108で所定の削除期間が経過すると、タイマ108はグルー論理113に削除信号を送り、グルー論理113はスケジューラ118に削除信号を送る。このとき、スケジューラ118はこの満了した削除期間に対応する(前にもとのポートに対応する待ち行列121に入れられた)待ち行列エントリを削除する。

【0101】待ち行列エントリが、スケジューラ118によって待ち行列121に入れられるとき、スケジューラ118は同時にアービトラータ123に特定のデータ結線と宛先ポート33に対する転送要求を送る。スケジューラ118はグルー論理113から次宛先ポート信号を受け取り、この次宛先ポート33に対応する待ち行列

の先頭の待ち行列エントリを検索する。スケジューラ118はこの検索された待ち行列エントリに対応してアービトラータ123に転送要求を送る。

【0102】アービトラータ123は、出所ポートおよび宛先ポートのインテリジェンス機構73からの状態信号(たとえば、rxready信号、混合ready信号、混合バスready信号およびtxready信号)を継続的にモニタし、これらの特定の出所ポートおよび宛先ポート33がデータのやりとりが可能な状態であるかどうかを判定する。アービトラータ123はこの転送がクラス1転送であるかクラス2転送であるかを判定する。クラス1である場合、アービトラータ123はこの特定の出所ポート33および宛先ポート33の両方がデータの送信と受信の両方を行なうことができる状態であるかどうかを判定する、これはクラス1転送には専用の転送路が必要であるためである。クラス2である場合、アービトラータ123はこの特定の出所ポート33からの転送が可能であるかどうか(すなわち出所ポート33についてrxready信号が出ているか)、またこの特定の宛先ポート33が新しいフレームを受信できるかどうか(すなわち宛先ポート33についてtxready信号が出ているか)を判定する。

【0103】ポート33が(クラス1あるいはクラス2の)転送を行なうことのできる状態ではないとき、アービトラータ123はスケジューラの転送要求を拒絶し、スケジューラ118は転送要求が最終的に認められるまで転送要求を繰り返す。

【0104】アービトラータ123がクラス1転送要求を認めると、アービトラータ123はMDN42あるいはIDN44がデータ転送路を確保し、適当な出所ポート33を適当な宛先ポート33に接続するようにMDN42あるいはIDN44を構成する。さらに、アービトラータ123がクラス2転送要求を認めると、アービトラータ123はこの新しいデータフレームがその出所ポート33に対応する出所受信メモリ84から適当な宛先ポート33に送られるようにMDN42あるいはIDN44を構成する。

【0105】クラス1データ転送あるいはクラス2データ転送のいずれかが認められると、シーケンサ101はこの宛先ポート33に対応する特定の状態/制御機構85のセットアップを可能にするためにこの機構にデータの着信が予想されることを知らせる。また、アービトラータ123は、スケジューラ118が対応する宛先待ち行列から関連の待ち行列エントリを削除するようにスケジューラ118にスケジューラクリア信号を送り、またタイマ108にタイマクリア信号を送ってこのフレームに対応するタイマの内部fbsy/削除クロックをクリアする。さらに、タイマ108がそのクロックをクリアした後、タイマ108は出所ポート33に対応する状態/制御論理85に制御信号を送って、受信メモリ84に

この新しい信号（あるいはフレーム拒絶信号）をスイッチモジュール 36 と宛先ポート 33 に送らせる。

【0106】クローザ 128 は、転送が発生したことを確認し、この転送に関する統計を収集することによってフレーム転送を終了する。クローザ 128 はシーケンサ 101 を介したこのフレームの転送の状態をシーケンサ 101 を介して要求する。これに対して、宛先ポートインテリジェンス機構 73 がクローザ 128 にこの状態を知らせる。この状態情報には少なくとも、CRC エラー、空フレーム、f b s y、拒絶フレーム（不良 S I D あるいは D I D）、S O F c l、混合、および出所ポート番号が含まれる。クローザ 128 は E C インターフェース 133 にカウント値をインクリメントするように指示し、E C インターフェース 133 は適当なカウント値をインクリメントする。

【0107】当業者には、転送路割り当てシステム 50 の実施例とその方法および光ファイバースイッチ 30 に対して本発明の原理から逸脱することなくさまざまな変更や修正を加えるものであることが理解されるであろう。かかる変更や修正はすべて特許請求の範囲に示す本発明の範囲に該当するものである。

【0108】以上の内容をまとめると、次のようになる。

1. 光ファイバネットワークにおいて光ファイバチャンネル (32) を選択的に相互接続する光ファイバースイッチ (30) を介した高性能のデータ転送のための転送路割り当てシステム (50) であって、前記チャンネル (32) のそれぞれに対応する複数のポート (33) と、出所ポート (33) から着信した新しいデータフレーム (11) を格納するための前記ポート (33) に関係付けられたメモリ手段 (84) と、前記ポート (33) を相互接続するように構成された分配ネットワーク (362) と、前記メモリ手段 (84) が前記新しいデータフレーム (11) を受け取ったことを判定するように構成された、前記メモリ手段 (84) につながったセンタリ (104) と、前記ポート (33) のそれぞれに対応する宛先待ち行列 (121) を維持するため、前記センタリ (104) が前記新しいデータフレーム (11) の受信を判定した後前記センタリ (104) から前記新しいデータフレーム (11) に関する、出所ポート (33) を同定し、前記メモリ手段 (84) 内における前記フレームの位置を同定するメモリアドレスを同定する出所ポート標識と宛先ポート (33) を同定する宛先ポート標識を含む転送路データを受け取るように構成され、また前記転送路データを前記宛先ポート (33) に対応する宛先待ち行列 (121) に入れ、前記宛先待ち行列 (121) から前記転送路データを検索するように構成されたスケジューラ (118) と、前記分配ネットワーク (36) を制御するように構成され、前記スケジューラ (118) と前記ポート (33) につながれ、前

記宛先ポート (33) が他のフレーム転送要求をサービスするためにビジーであること、および前記宛先ポート (33) が前記フレーム転送要求を受けることができることを判定するように構成され、また利用可能性に基づいて前記転送要求を拒絶および許可するように構成されたアービトラータ (123) と、データ (11) の出所としての前記出所ポート (33) への将来のアクセスの順序を決めるための前記出所ポート (33) に関係付けられ、前記順序は拒絶された転送要求に対応する少なくとも 1 つの宛先ポート (33) を含む出所公平性保証手段 (130) と、からなることを特徴とする転送路割り当てシステム (50)。

【0109】2. 前記出所公平性保証手段 (130) は出所ポート (33) を 1 つだけ同定して前記順序を決める次出所標識であることを特徴とする上記 1 に記載の転送路割り当てシステム (50)。

3. 前記アービトラータ (123) は前記転送要求はパケット転送を確立するための要求であるか専用転送路転送を確立するための要求であるかを判定するクラス判定手段を有し、前記専用転送路転送が確立されるとき、前記アービトラータ (123) は前記宛先ポート (33) がデータ (11) の受信と送信の両方に使用可能であるか、また前記出所ポート (33) がデータ (11) の送信に使用可能であるかどうかを判定することを特徴とする上記 1 項又は 2 に記載の転送路割り当てシステム (50)。

【0110】4. 前記データ (11) の宛先としての前記宛先ポート (33) への将来のアクセスの他の順序を決めるための前記出所ポート (33) に関係付けられた宛先公平性保証手段 (130') であって、前記他の順序は前記宛先ポート (33) への他の拒絶された転送要求に対応する少なくとも 1 つの出所ポート (33) を含む宛先公平性保証手段 (130') を有することを特徴とする上記 1 乃至 3 の何れかに記載の転送路割り当てシステム (50)。

【0111】5. 前記データ (11) の宛先としての前記出所ポート (33) への将来のアクセスの他の順序を決めるための前記出所ポート (33) に関係付けられ、前記他の順序は前記出所ポート (33) への他の拒絶された転送要求に対応する少なくとも 1 つの出所ポート (33) を含む宛先公平性保証手段 (130') を有することを特徴とする上記 1 乃至 4 の何れかに記載の転送路割り当てシステム (50)。

【0112】6. 前記データ (11) の出所としての前記宛先ポート (33) への将来のアクセスの他の順序を決めるための前記宛先ポート (33) に関係付けられ、前記他の順序は拒絶された転送要求に対応する少なくとも 1 つの宛先ポート (33) を含む宛先公平性保証手段 (130) を有することを特徴とする上記 1 乃至 5 の何れかに記載の転送路割り当てシステム (50)。

【0113】7. 前記スケジューラ(118)は前記宛先待ち行列(121)を順次サービスするように構成されたことを特徴とする上記1乃至6の何れかに記載の転送路割り当てシステム(50)。

【0114】8. 光ファイバーネットワークにおいて光ファイバーチャネル(32)を選択的に相互接続する光ファイバースイッチ(30)を介した高性能のデータ転送のための転送路割り当てシステム(50)であって、前記チャネル(32)のそれぞれに対応する複数のポート(33)と、出所ポート(33)から着信した新しいデータフレーム(11)を格納するための前記ポート(33)に関係付けられたメモリ手段(84)と、前記ポート(33)を相互接続するように構成された分配ネットワーク(362)と、前記メモリ手段(84)が前記新しいデータフレーム(11)を受け取ったことを判定するように構成された、前記メモリ手段(84)につながったセントリ(104)と、前記ポート(33)のそれぞれに対応する宛先待ち行列(121)を維持するため、前記セントリ(104)が前記新しいデータフレーム(11)の受信を判定した後前記セントリ(104)から前記新しいデータフレーム(11)に関する、出所ポート(33)を同定し、前記メモリ手段(84)内における前記フレームの位置を同定するメモリアドレスを同定する出所ポート標識と宛先ポート(33)を同定する宛先ポート標識を含む転送路データを受け取るように構成され、また前記転送路データを前記宛先ポート(33)に対応する宛先待ち行列(121)に入れ、前記宛先待ち行列(121)から前記転送路データを検索するように構成されたスケジューラ(118)と、前記分配ネットワーク(36)を制御するように構成され、前記スケジューラ(118)と前記ポート(33)につながれ、前記宛先ポート(33)が他のフレーム転送要求をサービスするためにビジーであること、および前記宛先ポート(33)が前記フレーム転送要求を受け取ることができることを判定するように構成され、また利用可能性に基づいて前記転送要求を拒絶および許可するように構成されたアービトラータ(123)と、データ(11)の宛先としての前記宛先ポート(33)への将来のアクセスの順序を格納するための前記宛先ポート(33)に関係付けられ、前記順序は拒絶された転送要求に対応する少なくとも1つの出所ポート(33)を含む宛先公平性保証手段(130')とからなることを特徴とする転送路割り当てシステム(50)。

【0115】9. 対応するポート(33)を介して光ファイバースイッチ(30)に接続された光ファイバーチャネル(32)を選択的に相互接続するための光ファイバースイッチ(30)を介したデータ(11)の効率的で信頼性の高い転送を保証する方法(130)であって、前記スイッチ(30)において出所ポート(33)からの宛先ポート(33)に切り換えるべきデータ(1

1)を受け取るステップと、前記宛先ポート(33)が前記データ(11)の受信に使用可能であること、および前記宛先ポート(33)が他のデータ(11)の受信のためにビジーであることを判定するステップと、前記出所ポート(33)が前記データ(11)の送信に使用可能であること、および前記出所ポート(33)が他のデータ(11)の送信のためにビジーであることを判定するステップと、前記出所および宛先ポート(33)が使用可能であるとき、前記出所ポート(33)から前記宛先ポート(33)に前記データ(11)を転送するステップと、前記出所および宛先ポート(33)のいずれかがビジーであるとき、前記出所ポート(33)から前記宛先ポート(33)に前記データ(11)の転送を拒絶するステップと、データ(11)の出所としての前記出所ポート(33)への将来のアクセスの順序を決めるため、前記順序は少なくとも1つの拒絶された転送要求に対応する少なくとも1つの宛先ポート(33)の指定を含むステップと、前記出所ポート(33)の前記ビジー状態が終わった後、前記順序に基づいて宛先ポート(33)の指定を獲得することによって拒絶された転送要求を可能にすることを特徴とする方法(130)。

【0116】10. 対応するポート(33)を介して光ファイバースイッチ(30)に接続された光ファイバーチャネル(32)を選択的に相互接続するための光ファイバースイッチ(30)を介したデータ(11)の効率的で信頼性の高い転送を保証する方法(130)であって、前記スイッチ(30)において出所ポート(33)からの宛先ポート(33)に切り換えるべきデータ(11)を受け取るステップと、前記宛先ポート(33)が前記データ(11)の受信に使用可能であること、および前記宛先ポート(33)が他のデータ(11)の受信のためにビジーであることを判定するステップと、前記出所ポート(33)が前記データ(11)の送信に使用可能であること、および前記出所ポート(33)が他のデータ(11)の送信のためにビジーであることを判定するステップと、前記出所および宛先ポート(33)が使用可能であるとき、前記出所ポート(33)から前記宛先ポート(33)に前記データ(11)を転送するステップと、前記出所および宛先ポート(33)のいずれかがビジーであるとき、前記出所ポート(33)から前記宛先ポート(33)に前記データ(11)の転送を拒絶するステップと、データ(11)の宛先としての前記宛先ポート(33)への将来のアクセスの順序を決めるステップであって、前記順序は少なくとも1つの拒絶された転送要求に対応する少なくとも1つの出所ポート(33)の指定を含むステップと、前記順序に基づいて出所ポートの指定を獲得することによって拒絶された転送要求を可能にすることを特徴とする方法(130)。

【0117】

【発明の効果】以上のように、本発明の転送路割り当てシステム(50)は、光ファイバーネットワークにおいて光ファイバーチャンネル(32)を選択的に相互接続する光ファイバースイッチ(30)を介した高性能のデータ転送のための転送路割り当てシステム(50)であって、前記チャンネル(32)のそれぞれに対応する複数のポート(33)と、出所ポート(33)から着信した新しいデータフレーム(11)を格納するための前記ポート(33)に関係付けられたメモリ手段(84)と、前記ポート(33)を相互接続するように構成された分配ネットワーク(362)と、前記メモリ手段(84)が前記新しいデータフレーム(11)を受け取ったことを判定するように構成された、前記メモリ手段(84)につながったセントリ(104)と、前記ポート(33)のそれぞれに対応する宛先待ち行列(121)を維持するため、前記セントリ(104)が前記新しいデータフレーム(11)の受信を判定した後前記セントリ(104)から前記新しいデータフレーム(11)に関する、出所ポート(33)を同定し、前記メモリ手段(84)内における前記フレームの位置を同定するメモリアドレスを同定する出所ポート標識と宛先ポート(33)を同定する宛先ポート標識を含む転送路データを受け取るように構成され、また前記転送路データを前記宛先ポート(33)に対応する宛先待ち行列(121)に入れ、前記宛先待ち行列(121)から前記転送路データを検索するように構成されたスケジューラ(118)と、前記分配ネットワーク(36)を制御するように構成され、前記スケジューラ(118)と前記ポート(33)につなぐれ、前記宛先ポート(33)が他のフレーム転送要求をサービスするためにビジーであること、および前記宛先ポート(33)が前記フレーム転送要求を受け取ることができることを判定するように構成され、また利用可能性に基づいて前記転送要求を拒絶および許可するように構成されたアービトラータ(123)と、データ(11)の出所としての前記出所ポート(33)への将来のアクセスの順序を決めるための前記出所ポート(33)に関係付けられ、前記順序は拒絶された転送要求に対応する少なくとも1つの宛先ポート(33)を含む出所公平性保証手段(130)と、からなるものである。

【0118】したがって、本発明によれば、高性能で、さまざまなアプリケーションに対するフレキシビリティがあり、転送要求がポートの利用可能である確率に依存することなく最終的にスイッチによって処理されることを保証でき、回路切り換えとフレーム切り換えの両方を行なうことができる。

【図面の簡単な説明】

【図1】光ファイバーネットワーク用の高性能光ファイバースイッチの概略回路図である。

【図2】図1の光ファイバースイッチのチャンネルモジ

ュール(CM)の概略回路図である。

【図3】図1の光ファイバースイッチ内の新しい転送路割り当てシステムの概略回路図である。

【図4】図3のアービトラータに採用された公平性保証アルゴリズムを示す概略図である。

【図5】図3のアービトラータのアーキテクチャと機能を示す高レベルフローチャートである。

【図6】図3のアービトラータのアーキテクチャと機能を図2の要素コントローラとの関連において示すフローチャートである。

【図7】図3のアービトラータのアーキテクチャと機能を図2の混合分配ネットワーク(IDN)との関連において示すフローチャートである。

【図8】図3のアービトラータのアーキテクチャと機能をクラス1(回路切り換え)データ転送ク(IDN)との関連において示すフローチャートである。

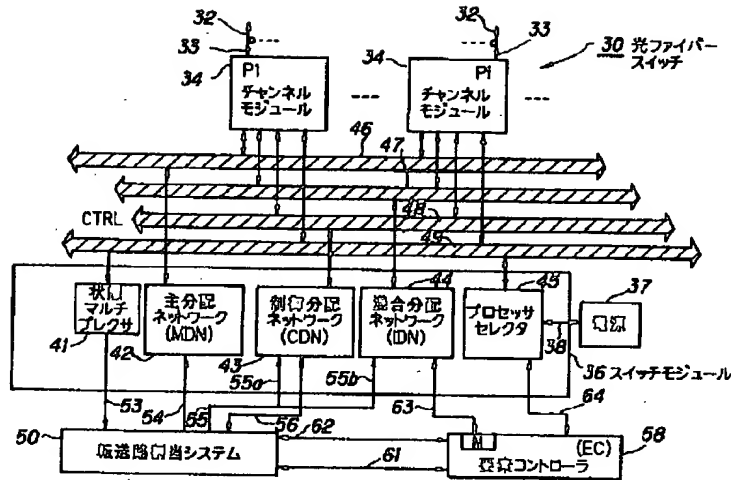
【図9】ファイバーチャンネル規格にしたがって光ファイバーネットワークの光ファイバースイッチを介して通信される可変長フレームの概略図である。

【符号の説明】

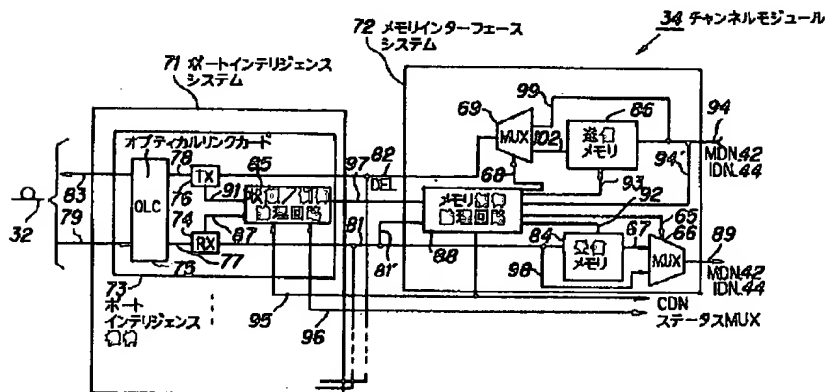
- 11: 可変長フレーム
- 12: フレーム始点(SOF)標識
- 14: ヘッダ
- 16: 可変長データ
- 17: CRC(周期冗長検査)符号
- 18: フレーム終点(Eof)標識
- 30: 光ファイバースイッチ
- 32: 光ファイバーチャンネル
- 33: ポート(p1~pi)
- 34: チャンネルモジュール
- 36: スイッチモジュール
- 37: 電源
- 41: 状態マルチプレクサ(MUX)
- 42: 主分配ネットワーク(MDN)
- 43: 制御分配ネットワーク(CDN)
- 44: 混合分配ネットワーク(IDN)
- 45: プロセッサセレクト
- 50: 転送路割り当てシステム
- 58: 要素コントローラ(EC)
- 61: データリンク
- 62: 状態/制御結線
- 64、116、117、122、126、134: 結線
- 72: メモリインターフェースシステム
- 73: ポートインテリジェンス機構
- 84: 受信メモリ
- 85: 状態/制御論理
- 86: 送信メモリ
- 101: シーケンサ
- 103、106、107、111、112、114、131: 制御結線

- | | |
|--------------------------------------|-----------------------------|
| 104: セントリ | 128: クローザ |
| 108: タイマ | 133: 要素コントローラ (EC) インターフェース |
| 113: グルー論理 | 135: シリアル結線 |
| 118: スケジューラ | 130: ネクスト・イン・ライン公平性保証アルゴリズム |
| 121: 宛先待ち行列 ($Q_{p1} \sim Q_{pi}$) | 05 ム |
| 123: アービトラータ | 130': 宛先公平性保証機構 |
| 124: データ結線 | 130'': 出所公平性保証機構 |
| 126: 結線 | 131: 待ち行列 Q_{p3} へのサービス |

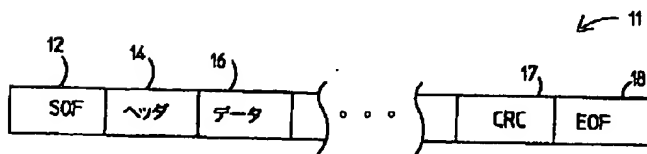
【図 1】



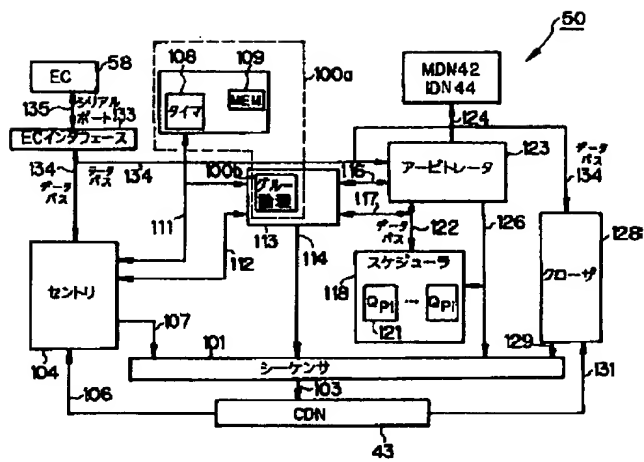
【図 2】



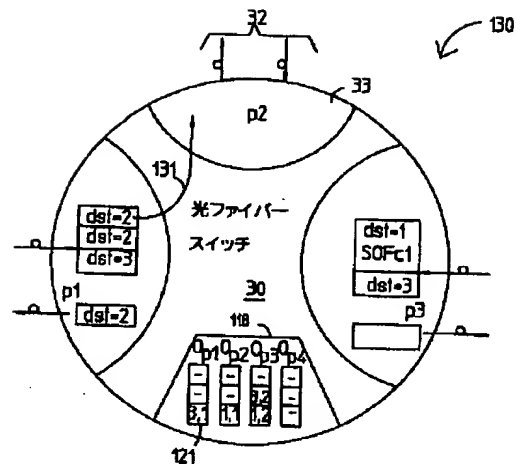
【図 9】



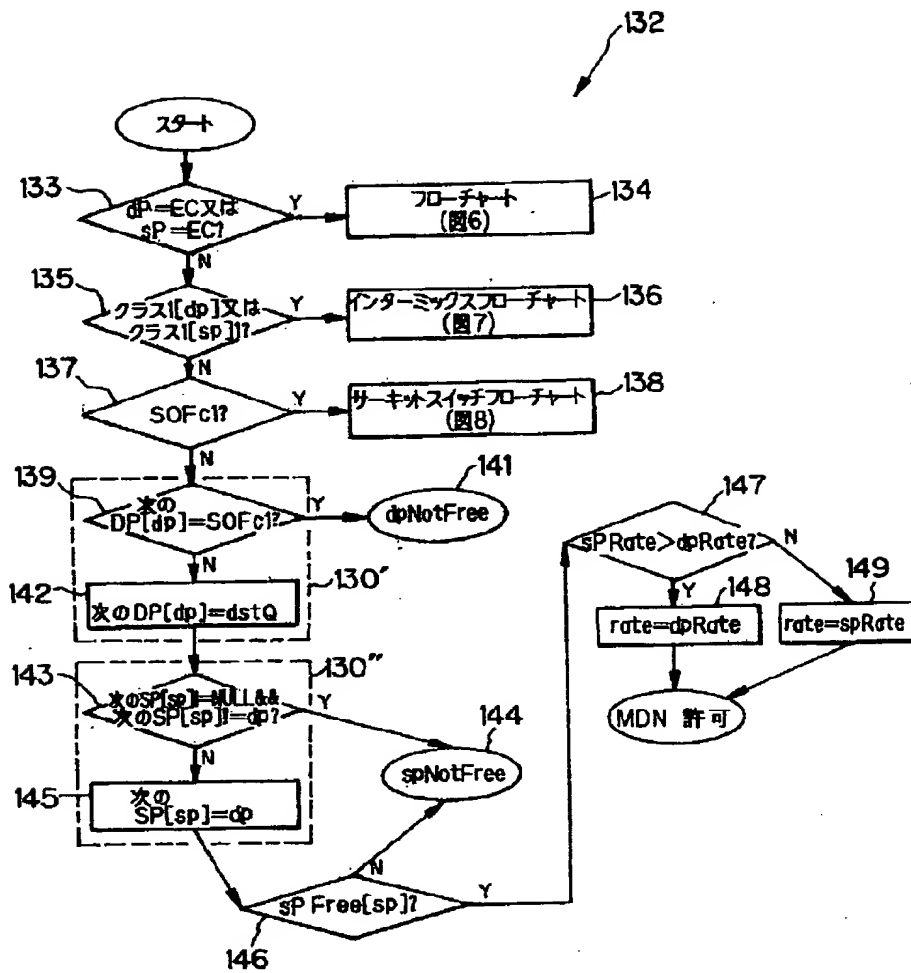
【図 3】



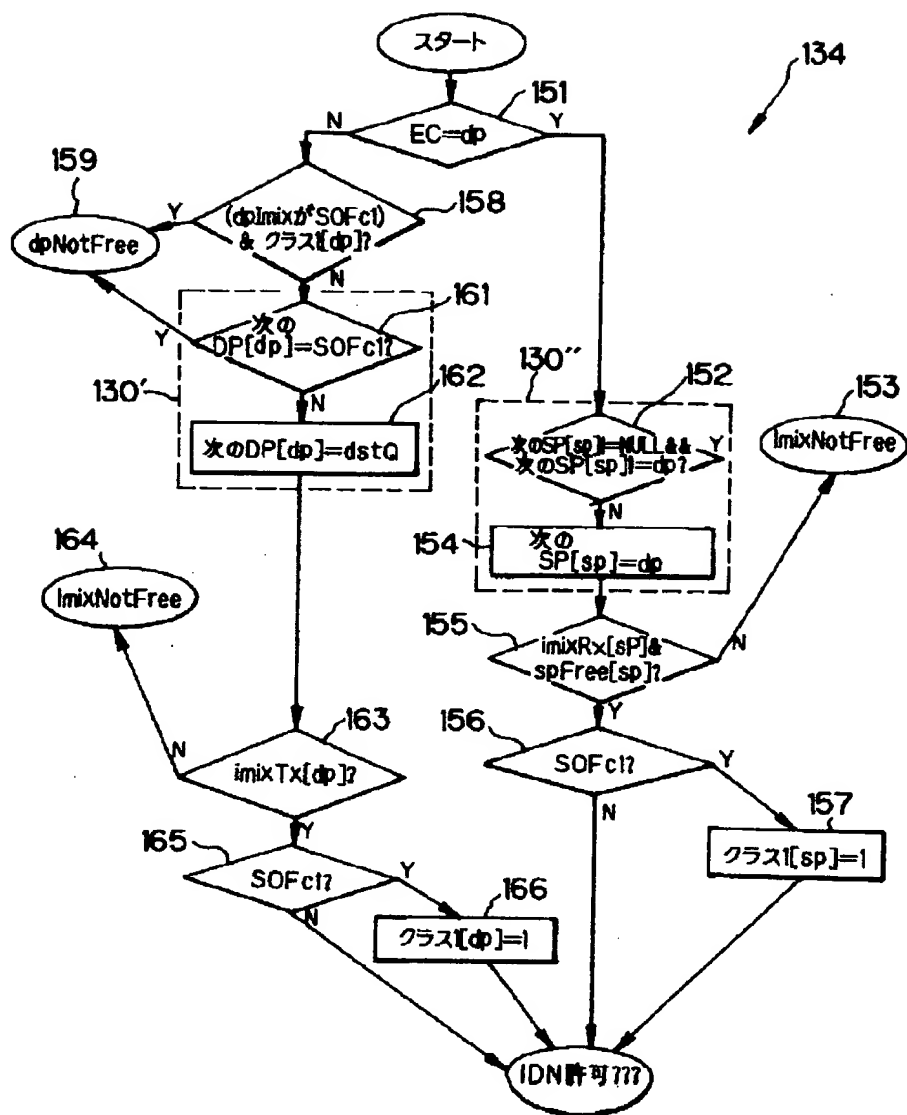
【図 4】



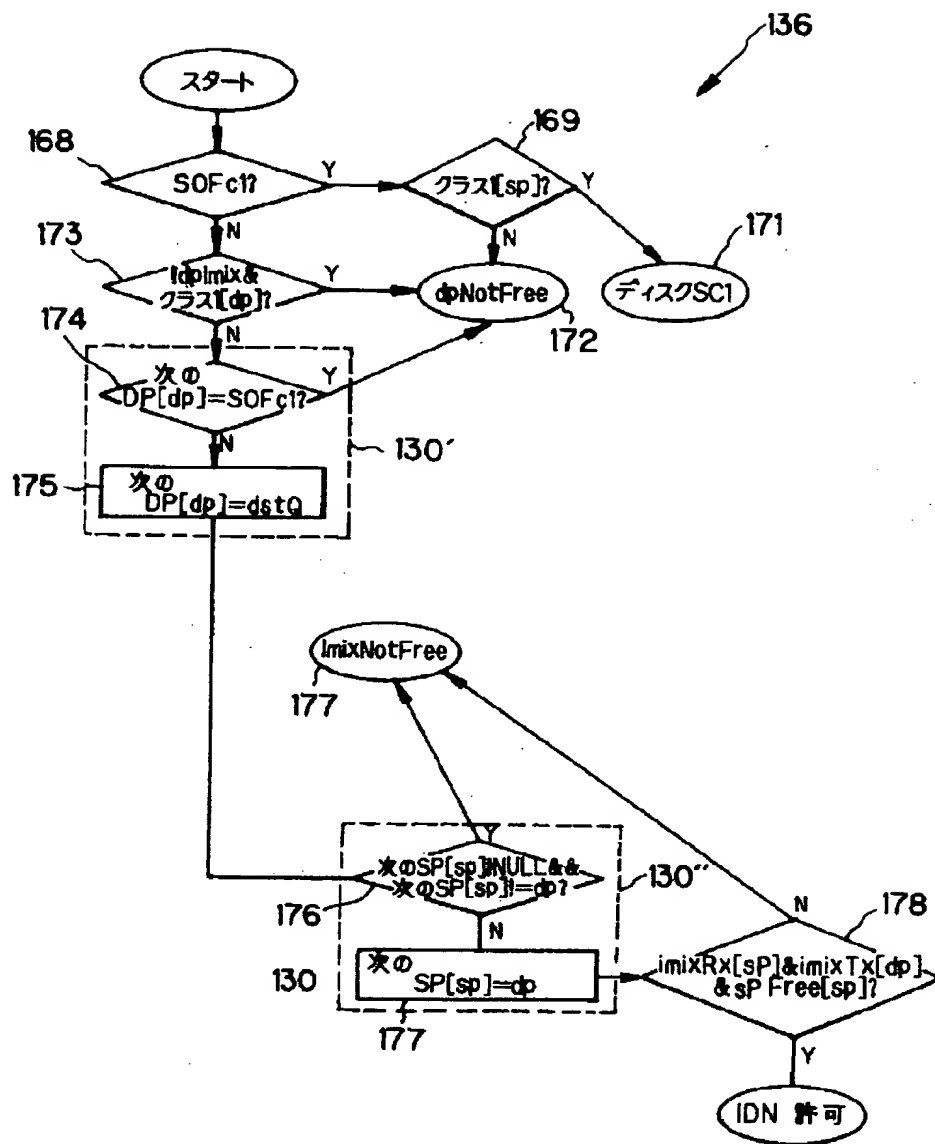
【図 5】



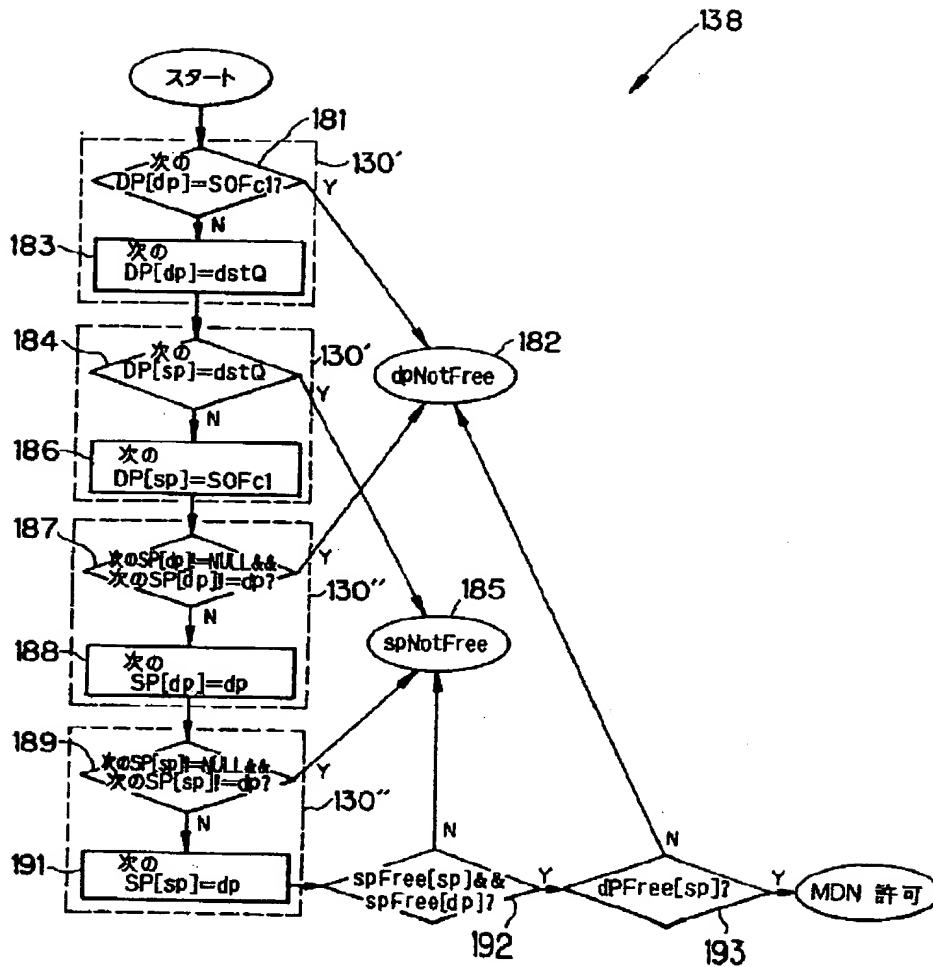
【図6】



【図 7】



【図 8】



フロントページの続き

(72)発明者 ロビン・ブルヒット
カナダ国オンタリオ州トロント メイジャ
ー・ストリート 193

40

(72)発明者 グレゴリー・ティー・サリヴァン
カナダ国オンタリオ州ブラントン シェン
ストーン・アヴェニュー 6

(72)発明者 デイヴィッド・ブック
カナダ国オンタリオ州ソーンヒル ヤン
グ・ストリート 7250、ユニット 511